CLA abundance estimates for western North Pacific Bryde's whales and their associated CVs with taking the additional variance into account

Toshihide Kitakado[1], Hiroyuki Shimada[2], Hiroshi Okamura[2] and Tomio Miyashita[2]

[1]*Tokyo University of Marine Science and Technology, Tokyo 108-8477, Japan*

[2]*National Research Institute of Far Sea Fisheries, Fisheries Research Agency, Kanagawa 236-8648, Japan*

The corresponding author's e-mail address: kitakado@kaiyodai.ac.jp

ABSTRACT

This paper aims at providing recent information on abundance of western North Pacific Bryde's whales, which could be used in the Catch Limit Algorithm (CLA). The estimation procedure consists of a total of three stages: 1) get sets of underlying abundance estimates in 1988-1996 and 1998-2002 surveys (dealt in Shimada et al., 2008), 2) estimate the process error (additional variance) in a mixed-effect model by a restricted maximum likelihood method based on the underlying abundance estimates, and then 3) provide point estimates for abundances in 1998-2002, by integrating the information from two different survey modes in 1998-2002 surveys, as well as their associated CVs by taking the extent of the process error into account. Note that abundance estimates from 1988-1996 surveys, which were not subject to oversight by IWC because of no oversight system at that time, did not affect the point estimates of abundances for CLA but was used only for the estimation of the process error. To assess the impact of model assumptions and data utilized on the estimation of the process error and abundances, computation were made under nine different run sets in total. Under the selected case in a base scenario, where the effective strip widths and mean school sizes were estimated by the sub-area and survey period and a log-normal mixed effect model was employed in the estimation of the process error, the estimated value of the additional CV was 0.447 (CV=43.1%). Taking the additional CV into account, the abundance estimates in sub-areas 1W, 1E and 2 were 4,957 (CV=39.8%), 11,213 (CV=49.8%) and 4,331 (CV=55.3%), respectively, and that in the whole area was 20,501 (CV=33.7%). The value of the additional CV under a count-based model was 0.535 (CV=50.2%) and was greater than the above value, which implied that the resultant CVs in abundance estimates were also higher than those in the log-normal model; CV=43.9%, 55.1%, 60.0% and 36.6% for sub-areas 1W, 1E and 2, and the whole area, respectively.

1. INTRODUCTION

The aim of this paper is to provide revised abundance estimates used in the Catch Limit Algorithm (CLA) for western North Pacific Bryde's whales. The estimation procedure consists of a total of three stages: 1) get sets of underlying abundance estimates in 1988-1996 and 1998-2002 surveys, 2) estimate the process errors (additional variance), and then 3) provide abundances estimates in 1998-2002 and their associated CVs by taking the extent of the process error into account. The present paper is the complement to our companion paper, Shimada et al. (2008), in which detailed information on surveys and analyses for underlying abundance estimation is summarized.

There are some modifications in the estimation procedure from our earlier paper. In Shimada et al. (2007), which was submitted to the Scientific Committee held in Anchorage, abundance estimates for sub-areas were provided as parametric estimates using integrated models, in which area-effects, year trend, a correction factor (between two different survey modes) and the additional variance were included as unknown parameters. However, in that model, abundance estimates from past surveys (1988-1996), which were not subject to oversight administrated by IWC, were used as dependent variables in addition to the recent surveys (1998-2002) overseen by IWC. The past data were needed because the recent surveys unfortunately covered only once in each surveyed blocks, and therefore it was not possible to estimate the additional variance unless additional information on abundance is available. However, such an estimation method might have turned out to draw information on abundance from surveys which were not subject to IWC oversight, and therefore the method virtually violated the regulation in RMP. To overcome this difficulty, we here take an alternative estimation procedure illustrated below. The additional variance is estimated using both the underlying abundance estimates in 1988-1996 and 1998-2002 surveys. Meanwhile, the point estimates of abundances in 1998-2002 in the last stage are estimated based only on the data from the recent surveys although their CVs depend on the value of the additional variance estimated with both the surveys. Under this procedure, although the variances of estimates still depend on abundance estimates in the past surveys, abundance estimates themselves do not depend on the information from the past surveys.

To assess the impact of model assumptions and data utilized on the estimates of the process error and CLA abundances, computation were made under totally nine run sets. These were prepared to respond to some of the constructive suggestions made during SC59 (IWC 2008; see Appendix). Although the estimation procedure in this present paper is slightly different from that in our earlier paper as described above, most of these suggestions were still effective. All the responses to the suggestions for the estimation on additional variance and their relevant matters were shown in Appendix.

2. MATERIALS AND METHODS

2.1 Underlying abundance estimates in surveyed blocks

The underlying abundance estimates are available from two surveys (past=1988-1996, recent=1998-2002; see Shimada et al. 2008). The recent survey was conducted with the two survey modes, “abeam closing mode” (closing starts after abeam of a school detected) and “normal closing mode” (closing starts after detection of a school), while the past survey was done by only the normal closing mode. The effective strip with (ESW) and mean school size (MSS) were estimated in several stratifications, and then the best set was selected (Shimada et al. 2008). Sampling variances for the abundance estimates in blocks were assessed by the conventional ways. However, the variance components for the encounter rates were estimated with leg-unit for the recent survey while those for the past survey have been estimated with 1-degree sector, and therefore there is inconsistency in sampling unit between past and recent surveys. To resolve this discrepancy in the procedure for the estimation uncertainty, a calibration factor was defined and used for adjusting CVs in Past period of survey (Shimada et al. 2008). Some of abundance estimates were not useful due to poor coverage or little

efforts in blocks. Abundance estimates used for this estimation were discussed in Pre-Implementation Assessment in 2005 (IWC 2006), and this decision has still been followed in this paper (except for an option that includes no sighting "0" data. See explanation on the run sets below). The information on abundances estimates and their CVs, observed school counts, and survey efforts used in this paper are summarized in Table 1. A map of the sub-areas and blocks is shown in Figure 1.

2.2 Estimation of the process error (additional variance)

The issue of the process error or additional variance arises from the fact that the estimated sampling variances for the abundance estimates do not account for variability of abundance level, especially due to inter-annual changes in distribution of whale population in the surveyed areas. If the additional variance is ignored, uncertainty on abundance estimates tends to be underestimated. Remark that we will use "process error" and "additional variance" interchangeably. Also, we here estimate the extent of the process error as the additional CV. The model specification will be explained later.

In order to estimate the amount of process error, replicates of surveys in each block in different years are usually required. However, the recent 1998-2002 survey covered each block only once and hence only the recent surveys themselves never provide such information on the process errors. For this purpose, we made use of the past abundance estimates as well as the recent abundance estimates.

For the estimation of the additional variance, log-normal models for the underlying abundance estimates have been conventionally used (Cooke 1994, Punt et al. 1997, Skaug et al. 2004 and Kitakado et al. 2005). Here, we employed a version of this model as a base model. The mean and variance in the log-normal distribution were given by keeping consistency in those moments. The additional variance model has usually area-effects as covariates. In addition, year factors are also taken into account in several manners. As described above, we have abundance estimates from the abeam and normal closing modes. The estimates based on the normal closing mode tend to lead to bias because breaking off the survey effort after the detection of school may cause skip of high density area. Hence, we also incorporate the survey mode as a covariate. The log-normal model with several covariates is used in integrated analyses to take the several factors into account. A count model is also used to deal with "0" data. More detailed explanation on the estimation procedure for process errors is given below. The extent of additional variance is estimated with a REML method. Better covariates in the model are investigated by AIC based on the likelihood (not the integrated likelihood for REML because the latter does not have information on covariates).

2.2.1 Statistical models and estimation: log-normal model

Suppose that an area consisting of A blocks is covered by each of two survey periods. Assume that each block is surveyed at least once in each of the survey periods. Let $\tilde{N}_{ay}$ be the actual abundance in the a-th block in year y, and let $\hat{N}_{ay}^{(AC)}$ and $\hat{N}_{ay}^{(NC)}$ be estimates of $\tilde{N}_{ay}$ by the abeam closing mode and the normal closing mode surveys, respectively. We assume that the abundance estimates are multivariate log-normally distributed

as follows:

$$\begin{cases} \log \hat{N}_{ay}^{(AC)} = \log \tilde{N}_{ay} + \varepsilon_{ay}^{(AC)}, \\ \log \hat{N}_{ay}^{(NC)} = \log R + \log \tilde{N}_{ay} + \varepsilon_{ay}^{(NC)}, \end{cases} \tag{1}$$

where $\varepsilon_{ay}^{(AC)}$ and $\varepsilon_{ay}^{(NC)}$ are the vectors of terms expressing the sampling errors in log-abundance, and R is a so-called correction factor. For notational convenience, we denote by $\hat{N}$, $\tilde{N}$ and ε the vectors of all the abundance estimates, actual abundances and sampling errors, respectively. The variance of abundance estimate (Buckland et al. 2001) is given by

$$\hat{\mathrm{var}}\left(\hat{N}_i\right) = \hat{N}_i^{\,2} \cdot \left\{ \frac{\mathrm{var}(n_i)}{n_i^{\,2}} + \frac{\mathrm{var}(\hat{w})}{\hat{w}^2} + \frac{\mathrm{var}\left(\hat{E}[s_i]\right)}{\hat{E}[s_i]^2} \right\} = \hat{N}_i^{\,2} CV_i^2. \tag{2}$$

If abundance estimates share common parameters such as the ESW or MSS across years, blocks and survey modes, these error terms are correlated. For example, if two abundance estimates, $\hat{N}_i$ and $\hat{N}_j$, share a common ESW and a common MSS, then the covariance between the two abundance estimates is give by

$$\hat{\mathrm{cov}}\left(\hat{N}_i, \hat{N}_j\right) = \hat{N}_i \hat{N}_j \cdot \left\{ \frac{\mathrm{var}(\hat{w})}{\hat{w}^2} + \frac{\mathrm{var}\left(\hat{E}[s]\right)}{\hat{E}[s]^2} \right\} = \hat{N}_i \hat{N}_j CV_{ij}^2 \ . \tag{3}$$

To make $\varepsilon = (\varepsilon_1, ..., \varepsilon_d)'$ satisfy the unbiasedness of abundance estimates and the variance-covariance formula above, we assume the mean, variance, and covariance of the errors for log-abundance as follows:

$$\begin{aligned} &E[\varepsilon_i] = -\frac{1}{2}\log(1 + CV_i^2)\ , \\ &Var[\varepsilon_i] = \log(1 + CV_i^2), \\ &Cov[\varepsilon_i, \varepsilon_j] = \log(1 + CV_{ij}^2) \quad (i \neq j). \end{aligned} \tag{4}$$

We now assume that the true abundance level varies randomly over years as

$$\log \tilde{N}_{ay} = \log N_{ay} + \rho_{ay}\,, \tag{5}$$

where N_{ay} is an expected abundance in the *a*-th block in year *y*, and ρ_{ay} is a random effect accounting for inter-annual changes in the distribution of the whale population in the surveyed area. The random effects are assumed to be independent and identically distributed as the normal distribution,

$$\rho_{ay} \sim N(-\frac{1}{2}\log(1+\sigma^2), \log(1+\sigma^2))\,, \tag{6}$$

where σ can be called the "additional coefficient of variation" since $Var[\tilde{N}_{ay}] = \sigma^2 N_{ay}^2$ holds by (5) and

(6).

Next, we consider models for the expected abundance level. Let μ_a is a mean area-specific log-abundance at a specific year or period (depends on models) in the *a*-th block. Then, we investigate the four models for N_{ay} as follows:

Case 1: True abundance level is constant over 1988-2002

$$\log N_{ay} = \mu_a$$

Case 2: True abundance level is exponentially increasing (deceasing) since 1988

$$\log N_{ay} = \mu_a + \alpha(2000 - y)$$

Case 3: True abundance levels are different between the two survey periods

$$\log N_{ay} = \mu_a, \quad for \quad y = 1998,...,2002$$
$$\log N_{ay} = \alpha + \mu_a, \quad for \quad y = 1988,...,1996$$

Case 4: Abundance levels differ in the two survey periods with the interaction between the period and Latitude (low, middle and high latitudes)

$$\log N_{ay} = \mu_a, \quad for \quad y = 1998,...,2002$$
$$\log N_{ay} = \alpha + \mu_a, \quad for \quad y = 1988,...,1996, \text{ and block } a \text{ is in the low latitude}$$
$$\log N_{ay} = \alpha + \mu_a + \beta_1, \quad for \quad y = 1988,...,1996, \text{ and block } a \text{ is in the middel latitude}$$
$$\log N_{ay} = \alpha + \mu_a + \beta_2, \quad for \quad y = 1988,...,1996, \text{ and block } a \text{ is in the high latitude}$$

Case5: Abundance levels differ in the two survey periods with the interaction between the period and block

$$\log N_{ay} = \nu_a, \quad for \quad y = 1988,...,1996$$
$$\log N_{ay} = \mu_a, \quad for \quad y = 1998,...,2002$$

Now the parameter of interest is the additional variance σ^2. It is well-known that the conventional ML method causes underestimation of the variance factor, and therefore we use an REML method for this purpose. However, the integrated likelihood for REML does not have information on covariates. Hence, model selection among the Cases above in the estimation of additional variance is done by AIC based on the likelihood (not integrated with respect to effect parameters).

As in a conventional notation of mixed effect models, we represent the model as a linear form as follows:

$$
\begin{aligned}
&Y = X\beta + Z\rho + \varepsilon, \\
&\rho \sim N(-\frac{1}{2}diag(D), D), \quad D = \log(1+\sigma^2)\mathbf{I}, \\
&\varepsilon \sim N(-\frac{1}{2}diag(\hat{\Sigma})), \hat{\Sigma}), \quad \hat{\Sigma} = (\log(1+CV_{ij}^2))_{i,j}
\end{aligned}
\tag{8}
$$

Where $Y = \log \hat{N}$ is the vector of log-abundance estimates, X is a design matrix for the fixed-effects in linear predictor for $\log N_{ay}$, Z is a design matrix for random effects (to handle with the cases where the same area was surveyed in a year with different modes in the recent period), and D and $\hat{\Sigma}$ are the variance-covariance matrix for $\rho = (\cdots, \rho_{ay}, \cdots)'$ and ε.

For fixed σ^2, the best linear unbiased estimator of β is derived by

$$
\beta(\sigma^2) = (X'V(\sigma^2)^{-1}X)^{-1}X'V(\sigma^2)^{-1}Y^*, \tag{9}
$$

where $Y^* = Y + \frac{1}{2}Z\,diag(D) + \frac{1}{2}diag(\hat{\Sigma})$ and $V(\sigma^2) = ZDZ' + \hat{\Sigma}$.

The additional variance σ^2 is estimated by the REML method (Punt et al., 1997, McCulloch and Searle, 2001; Pawitan, 2001), which maximizes

$$
l_{REML}(\sigma^2) = -\frac{1}{2}\log\left|V(\sigma^2)\right| - \frac{1}{2}\log\left|X'V(\sigma^2)^{-1}X\right| - \frac{1}{2}(Y^* - X\beta(\sigma^2))'V(\sigma^2)^{-1}(Y^* - X\beta(\sigma^2)), \tag{10}
$$

and its uncertainty is assessed by the reciprocal of the second derivative with respect to σ^2.

2.2.2 Statistical models and estimation: negative binomial count model

The log-normal model described above can not be employed in the presence of “0” abundance estimates unless excluding them. To assess the impact of a way ignoring “0” abundance estimates in the estimation of additional variance, we also consider a count data model, where “0” data can be used..

Let n_i and m_i be the number of schools detected during a survey in a block in a year and its expected value, respectively. We then consider a model for the count n_i as

$$
\begin{aligned}
&n_i \sim NB(m_i, \lambda), \\
&\log m_i = \log(2L_i / A_i) + \log(\hat{w} / \hat{E}[s]) + u_i + \log \tilde{N}_i, \\
&\log \tilde{N}_i = \log N_i + \rho_i, \qquad \rho_i \sim (iid)\, N(-0.5\log(1+\sigma^2), \log(1+\sigma^2)),
\end{aligned}
\tag{11}
$$

where $NB(m_i, \lambda)$ shows a negative binomial distribution having the probability function

$$P(n_i) = \frac{\Gamma(n_i + \lambda)}{\Gamma(n_i + 1)\Gamma(\lambda)} \left(\frac{m_i}{m_i + \lambda} \right)^{n_i} \left(\frac{\lambda}{m_i + \lambda} \right)^{\lambda} \qquad (12)$$

with the mean $E[n_i] = m_i$ and variance $Var[n_i] = m_i(1 + m_i / \lambda)$, and L_i and A_i are the effort and the area, respectively. The error terms u_i's have normal distributions with known parameters to account for estimation uncertainty for the log-ratio of ESW to MSS, and if counts in blocks are associated with same ESW and MSS, then these error terms are definitely same random variable. N_i's are modeled with covariates as in Cases 1-5.

The REML function for dispersion parameters, σ^2 and λ, is defined as

$$L(\sigma^2, \lambda) = \int \cdots \int L(\beta, \lambda, \rho, u, \sigma^2) d\beta \, d\rho \, du, \qquad (13)$$

where $L(\beta, \lambda, \rho, u, \sigma^2)$ is the product of likelihood functions of the count data, random effects, and uncertainty in ESW/MSS. Although this function cannot be explicitly expressed in a closed form, a Laplace approximation numerically provides REML estimates of σ^2 and λ.

2.2.3 Specification of run sets for the estimation of the process errors

In this paper, impacts of assumptions on the estimation of additional variance are examined through a total of nine runs. These runs originally came up from the discussion at SC59. Although we altered the estimation procedure as described above, these kinds of investigation are still quite important. For example, we use, as base estimates, the abundances with ESWs and MSSs estimated by period and sub-area (see Run#1). As a sensitivity test, we also use another set of underlying abundance estimates with ESWs and MSSs estimated by only period (Run#2). It is expected that the correction factor for adjusting abundance estimates by the normal closing mode is less than 1, and it is possible to let the data tell if it is the case or not. But, in Run#3, we assumes R=1 instead of use of its estimate to know sensitivity to separation of survey modes (Run#7 is also related to correction factor). Run#4 is prepared to compare CVs in abundance estimates with/without consideration of additional variance. Run#5 is for generalization of regression models to allow interactions between blocks and period. Run#6 is to estimate the additional variance using abundance estimated only by the normal closing mode. This does not require any correction between the abeam and normal closing mode surveys. Run#8 aims at investigating the possibility to estimate additional variance based only on Past surveys. There are some replicates of surveys in some blocks during Past surveys, and hence it may be able to draw information on the additional variance. The additional variance is also estimated based on a count-based model to examine impact of ignoring "0" abundance estimates. Thus, a negative binomial model is used in Run#9. A more detailed description on the specification of the run sets is shown in Table 2.

2.3 CLA abundance estimates in 1998-2002 period

CLA abundance estimates of blocks, Sub-areas (IW, IE, and II), and the whole management Area were

provided. They were based only on estimates from the recent surveys (overseen). Abundance estimates from two different modes were combined using a correction factor, which was re-estimated based only on the data from the recent surveys. Furthermore, the additional variance was taken into account in their associated CVs. This is a different way from that in Shimada et al.(2007), where posterior means of "block-effects" in REML in the integrated model for the estimation of additional variance were used for abundance estimates for blocks although these depends on abundance estimates of the past surveys. The reason why we take the procedure below is not to use the abundance estimates in the past surveys, which were not overseen by IWC. Therefore, CLA abundance estimates in 1998-2002 depend on the data from Past surveys only in terms of their CVs. Statistical models and estimation procedures are described below.

While the additional variance is estimated with the data from both the survey periods, abundance in the recent period themselves are estimated by only by making use of data from the recent survey overseen by IWC. For this purpose, we use a log-linear model again as follows:

$$\begin{cases} \log \hat{N}_a^{(AC)} = \log N_a + \rho_a + \varepsilon_a^{(AC)}, \\ \log \hat{N}_a^{(NC)} = \log R^* + \log N_a + \rho_a + \varepsilon_a^{(NC)}. \end{cases} \tag{14}$$

Note that the meaning of the correction factor R^* is same as that in the previously illustrated integrated model but it is re-estimated with only the abundance estimates from the recent surveys.

These are re-expressed again as a linear form as

$$\begin{aligned} &Y = X\beta + Z\rho + \varepsilon, \\ &\rho \sim N(-\frac{1}{2} diag(\hat{D}), \hat{D}), \qquad \hat{D} = \log(1+\hat{\sigma}^2)\mathbf{I}\ , \\ &\varepsilon \sim N(-\frac{1}{2} diag(\hat{\Sigma})), \hat{\Sigma}), \qquad \hat{\Sigma} = (\log(1+CV_{ij}^2))_{i,j} \end{aligned} \tag{15}$$

where β is a vector of the logarithm of abundances in blocks and correction factor, and all the other vectors and matrices are only for the recent period. Note that the additional CV, $\hat{\sigma}$, was estimated by the method above.

The best unbiased estimator of β is then given by

$$\hat{\beta} = (X'V(\hat{\sigma}^2)^{-1}X)^{-1}X'V(\hat{\sigma}^2)^{-1}\left[Y + \frac{1}{2}Z\, diag(\hat{D}) + \frac{1}{2} diag(\hat{\Sigma})\right] \tag{16}$$

and its variance-covariance matrix is estimated as

$$C\hat{o}v(\hat{\beta}) = (X'\hat{V}^{-1}(\hat{\sigma}^2)X)^{-1}, \tag{17}$$

where $\hat{V}(\hat{\sigma}^2) = Z\hat{D}Z' + \hat{\Sigma}$.

Then, an unbiased estimator of $\eta = e^{\beta} = (N_1, ..., N_A, R^*)'$, which is a parameter vector of interest, is derived by exponential transformation of $\hat{\beta}$ with adjustment due to the log-normality of $\hat{\beta}$ as

$$\hat{\eta} = \exp\left\{\hat{\beta} - \frac{1}{2} diag(C\hat{o}v(\hat{\beta}))\right\} \qquad (18)$$

and its variance-covariance matrix is given by

$$C\hat{o}v(\hat{\eta}) = \hat{P}\left\{\exp(C\hat{o}v(\hat{\beta})) - E\right\}\hat{P}, \qquad (19)$$

where $\hat{P}$ is a diagonal matrix with diagonal components $\hat{\eta}$, and E is a $(A+1)\times(A+1)$ matrix with all the elements equal to 1.

This procedure can be regarded as a sort of extension of Haw's method (Haw 1991) when correlations between abundance estimates are taken into consideration. It should be noted that the vector of estimates $\hat{\eta}$ is unbiased and does not depend on the value of $\hat{\sigma}^2$, which is estimated using abundance estimates from both the periods. Essentially, the correction factor R^* is estimated through the information on the differences of abundance estimates with the different modes within blocks, $\log \hat{N}_a^{(AC)} - \log \hat{N}_a^{(NC)}$, where the terms of random effects are cancelled out, and therefore R^* does not depend on $\hat{\sigma}^2$. Furthermore, although each block effect in log-scale depends on the value of $\hat{\sigma}^2$ due to the adjustment of the mean in $\rho \sim N(-0.5diag(\hat{D}), \hat{D})$, the exponential transformation with the adjustment in (18) also cancelled out such the effects. Hence, the abundance estimates in (18) do not depend on the value of the additional variance. Remark that $\hat{\eta}$ becomes the vector of the nominal abundance estimates $(\hat{N}_1, ..., \hat{N}_A)'$ when only the abundance estimates by the abeam closing survey mode are used in the procedure, and then the variance-covariance matrix is also same as the nominal one if the random effects are ignored (i.e. $\hat{\sigma}^2 = 0$).

3. RESULTS AND DISCUSSION

Table 3 shows results for Run#1 (the base cases, where the selected set of ESWs and MSSs was used in the underlying abundance estimates, Cases 1-4), Run#4(a) (No additional variance) and Run#5(a) (Case 5). Only the estimates of parameters of interest are provided in the table. AIC model selection indicated that Case 4 (block-effect and period-latitude interaction) was the best model for the additional variance model. A notable thing is that the estimates of R and R^* under Case 4 were almost same. This also imply that Case 4 in the additional variance model caught the actual distributional change to the north in Recent period (α=1.50, β_1=-3,47, β_2=-2.95). The observed and predicted log-abundances were plotted in Figure 2-(i). These graphs showed that the fitness in Cases 4 and 5 was better than that in other three cases although Case 5 seemed to

use parameters overly. The additional CV was estimated as 44.7%. This estimate was the smallest among the values in this runs though it indicated that a large amount of random distributional change occurred in the scale of blocks defined in this analysis. The CLA abundance estimates for the three sub-areas as well as the total area were also shown in the table. The incorporation of the additional variance increased more than about 30% of CVs compared to those with no additional variance.

Table 4 shows results for Run#2 (the base cases, where the secondary selected set of ESWs and MSSs was used in the underlying abundance estimates, Cases 1-4), Run#4(b) (No additional variance) and Run#5(b) (Case 5). This table also suggested that Case 4 was the best model for the additional variance model although the difference in AIC between Cases 4 and 5 was smaller than that in Run#1. This may be because the assumption of the common ESW and MSS masked the actual behaviors of abundance level. The estimates of the additional CVs were slightly different between Run#1 and Run#2 although those of R and R^* were quite similar. A possible reason is that the CVs for the underlying abundance in Run#2 (ESW and MSS were common to sub-area) were originally smaller and therefore the total variances of the abundance estimates were more explained by the additional variance in Run#2 than in Run#1. This makes sense since the total variance in abundance estimates is essentially decomposed into of the sampling variance and the additional variance as described in the model. The difference of resultant CVs in CLA abundance estimates between Run#1 and Run#2 were about 10% for the sub-area and the whole area except for the case of no additional variance.

Table 5 shows results of further sensitivity tests. Run#3 was conducted under the model of Case 4 to know sensitivity to separation of survey modes. The estimate of the additional CV was similar with that in Run#1 although the values of the correction factor were different. This result showed robustness of the additional CV to the correction factor. Remark that, in Run#3 and Run#7, the estimation uncertainty of the correction factor was not taken into account in CVs in the final abundance estimates, and therefore their CVs were not comparable with those in other runs. In Run#6, which aimed at estimating the additional variance using data from only normal closing mode, the additional CV was estimated slightly greater than that in Run#1. This was the case for Run#8, where only abundance estimates from the past survey were employed. These results might have attributed to the fact that the numbers of data utilized in Run#6 and Run#8 were smaller than that in Run#1. This could have caused less precise estimates of the additional CV, which was reflected the larger standard errors (not CVs) in the estimates in Run#6 and Run#8.

The results for Run#9 were shown in Table 6 and Figure 2-(ii). In this run, instead of assuming the log-normality of abundance estimates, negative binomial distributions were supposed for the block-wise counts of detected schools. This aimed at assessing the impact of ignoring the data with no sighting, which were eliminated from the data in runs other than Run#9. In this run, however, “0” data in the three blocks (1WW-M, 1WM-H, 1WM-M) were still eliminated because no sighting was made in those blocks through the whole two periods, and hence no information on the additional variance was drawn from those data. This run also selected Case 4 as the covariate for the additional variance model. Since the estimate of the additional CV in Case 4 was slightly larger than that in Run#1 (base case), CVs in the CLA abundance estimates were also slightly greater than those in the base case. In this analysis, counts were defined based on blocks because

leg-based efforts were not available in the past surveys (Shimada et al. 2008). The analysis by leg-based counts would be preferable to capture the variability in the counts between legs, and use of leg-based counts could decrease the difference in values of additional CV between Run#1 and Run#9.

Throughout these nine runs, the additional CVs were estimated around 45% or greater, and these estimates resulted in the high uncertainty in the abundance estimates. However, these high CVs were partially explained by an estimate in block 1E-H in 1993 survey, which took considerably large value. If this abundance estimate was eliminated from the analyses in Case 4 in Run#1, the resultant additional CV was 0.334 (CV=54.8%). Hence, only this estimate caused 34% of greater additional CV compared to the one when eliminating this abundance estimate. In this case, CVs decreased from 39.8%, 49.8% and 55.3% to 35.1%, 43.9% and 49.9% for sub-area 1W, 1E and 2, respectively, and from 33.7% to 30.3% for the whole area. Of course, we do not have any reasons to exclude that estimate, while this demonstrated sensitivity in the estimation of additional variance to the data set utilized.

Another sensitivity test (results not shown here) was conducted using underlying abundance estimates without making any correction of CVs in the past surveys to see the impact of the calibration. The calibration tended to increase the CVs in encounters (n/L) unless the number of detected schools was less than 4 (Shimada et al. 2008), and this treatment was expected to decrease the estimate of additional CV because of the trade-off between sampling and process errors. In fact, the resultant estimates with no calibration were 0.660, 0.668, 0.665, 0.521 and 0.664 for Cases 1-5, respectively. In Case 4, the case with the best covariates in our analyses, the value of 0.521 was greater than the value of 0.447 in Run#1 based on the corrected CVs while it was consistent with the value of 0.535 in Run#9, where no sampling CVs on the encounter rates were employed. In this sense, the results seemed to make sense.

We have shown a number of results under the several run sets. We would consider the result in Case 4 in Run#1 the best, but the result in Case 4 in Run#9 could also be worth referring to the best. The latter used the information on all the counts data except for those in blocks where no sighting have been made in the two survey periods. As results, in Case 4 in Run#1, where the effective strip widths and mean school sizes were estimated by the sub-area and survey period and a log-normal mixed effect model was employed in the estimation of the process error, the estimated value of the additional coefficient of variation was 0.447 (CV=43.1%). The abundance estimates in sub-areas 1W, 1E and 2 were 4,957 (CV=39.8%), 11,213 (CV=49.8%) and 4,331 (CV=55.3%), respectively, and that in the whole area was 20,501 (CV=33.7%). The value of the additional coefficient of variation under a count-based model in Case 4 in Run#9 was 0.535 (CV=50.2%) and was greater than the above value, which implied that the resultant CVs in abundance estimates were also higher than those in the log-normal model; CV=43.9%, 55.1%, 60.0% and 36.6% for sub-areas 1W, 1E and 2, and the whole area, respectively. The correlation matrices for the abundance estimates in blocks and the correction factor in these two results were shown in Table 7. The positive correlations between abundance estimates were made through the common ESWs and MSSs in the same sub-areas as well as the common correction factor (R^*). The negative correlations between the abundance estimates and R^* were consistent with the fact that larger value of R^* leads to smaller values of the abundance estimates. For reference to illustrate how the additional CV gives an impact on the uncertainty in abundances,

graphs for CVs in abundance estimates for different values of the additional CV were also shown in Figure 3.

The analyses here were based on blocks defined in the recent surveys. In the recent surveys, multiple blocks (2 or 3) were simultaneously covered in each of five years (i.e.; 1WW, 1WM, 1WE, 1E and 2), and therefore the analyses could have been done based on these five strata. However, this was not the case in the past survey. If those strata were applied for the estimation of the additional variance, less numbers of abundance estimates could have been used. This is a matter of planning of surveys, but this warrant that the future survey, if it will be conducted, should be planned to cover, in a single year, the blocks surveyed in a single year in the 1998-2002 surveys.

ACKNOWLEDGEMENT

The authors are grateful to Doug S. Butterworth and André E. Punt for their helpful comments. Also, the authors thank to Hans J. Skaug for his inspiring the count data modeling with ADMB-RE.

REFERENCES

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. and Thomas, L. 2001. *Introduction to Distance Sampling: Estimating abundance of biological populations.* Oxford University Press, New York, 432pp.

Cooke, J.G. 1994. Analysis of inter-survey variability of IDCR minke whale abundance estimates (`Process error').*Rep. Int. Whal. Commn.* 44:90.

Haw, M.D. 1991. An investigation into the differences in minke whale school density estimates from passing mode and closing mode surveys in IDCR Antarctic assessment cruises. Rep. int. Whal. Commn 41: 313–30.

IWC. 2006. Report of the Workshop on the pre-implementation assessment of western North Pacific Bryde's Whales. J. Cetacean Res. Manage. (Suppl.) 8.

IWC. 2008. Report of the Scientific Committee. Annex D. Report of the sub-committee on the Revised Management Procedure. J. Cetacean Res. Manage. (Suppl.) 10. In Press.

Kitakado, T., Shimada, H., Okamura, H. and Miyashita, T. 2005, Review of abundance estimate and additional variance for the western North Pacific stock of Bryde's whales, Paper SC/57/PFI1. 20pp.

McCulloch, C.E. and Searle, S.R. 2001. *Generalized, Linear, and Mixed Models.* Wiley & Sons, Inc, New York.

Pawitan, Y. 2001. *In All Likelihood: Statistical Modelling and Inference Using Likelihood.* Oxford University Press, New York.

Punt, A.E., Cooke, J.G., Borchers D.L. and Strindberg, S. 1997. Estimating the extent of additional variance for southern hemisphere minke whales from the results of the IWC/IDCR cruises. *Rep. Int. Whal. Commn.* **47**: 431-434.

Shimada, H., Kitakado, T., Okamura, H. and Miyashita, T. 2007. Abundance estimate of the western North Pacific stock of Bryde's whales using sighting data from 1998 to 2002. Paper SC/59/PFI3 presented to the IWC Scientific Committee, 2007 (unpublished).

Shimada, H., Okamura, H., Kitakado, T. and Miyashita, T. 2008. Abundance estimate of western North Pacific Bryde's whales for the estimation of additional variance and CLA application. Paper SC/60/PFI2 presented to the IWC Scientific Committee, 2008 (unpublished).

Skaug, H.J., Øien, N., Schweder, T. and Bølthun, G. 2004, Abundance of minke whales (*Balaenoptera acutorostrata*) in the Northeast Atlantic: variability in time and space, Can. J. Fish. Aquat. Sci. **61**: 870-886.

Table 1. Summary of underlying abundance estimates and their CVs in parentheses, school counts and survey efforts.

Blocks 1WW-M, 1WM-H and 1WM-M were excluded from the all the analyses because no detections have been made throughout the years.

(a) Abundance estimates (upper) and CVs (lower, %) under the assumption that both the ESW and MSS were stratified by the sub-area and survey period (the best case)

Year/Block	1WW-M	1WW-L	1WM-H	1WM-M	1WM-L	1WE-H	1WE-M	1WE-L	1E-H	1E-M	1E-L	2-H	2-M
Latitude	25°N-35°	10°N-25°N	35°N-43°N	25°N-35°N	10°N-25°N	31°N-43°N	20°N-31°N	10°N-20°N	32°N-43°N	20°N-32°N	10°N-20°N	35°N-43°N	25°N-35°N
Longitude	130°B-137°30'E	130°B-137°30'E	137°30'E-145°E	137°30'E-145°E	137°30'E-145°E	145°E-165°E	145°E-165°E	145°E-165°E	165°E-180°	165°E-180°	165°E-180°	180°-155W	180°-155W
Area (nm^2)	188,870	385,344	81,960	233,552	384,389	692,020	663,632	692,984	309,600	581,514	519,738	289,250	778,545
Past surveys (Normal closing)													
1988						**1,369** 52.96			**1,006** 57.83			**1,589** 59.53	
1989						**1,929** 47.85	**1,012** 50.67		**4,337** 57.87	**1,454** 65.84		**306** 78.28	**1,559** 90.30
1990						**1,100** 76.37	**806** 55.73		**1,510** 46.52	**2,854** 52.53		**398** 72.00	**757** 120.37
1991			**0**										
1992	**0**			**0**		**2,131** 44.97							
1993		**2,132** 58.12			**792** 56.27	**3,531** 128.05	**3,450** 53.48	**3,002** 71.14	**13,634** 74.27	**7,132** 130.87	**622** 74.28		
1994													
1995											**0**		
1996								**3,470** 45.06					
Recent surveys (Abeam closing)													
1998						**3,635** 73.06	**1,624** 52.90	**160** 111.48					
1999									**3,186** 56.63	**8,889** 73.56	**289** 110.70		
2000			**0**	**0**	**222** 105.74								
2001	**0**	**348** 88.08											
2002												**1,656** 56.49	**1,752** 83.96
Recent surveys (Normal closing)													
1998						**705** 40.02	**3,026** 48.71	**449** 57.43					
1999									**3,128** 35.00	**5,580** 49.58	**253** 63.66		
2000			**0**	**0**	**474** 70.79								
2001	**0**	**0**											
2002												**1,598** 74.31	**2,433** 55.25

(b) Abundance estimates (upper) and CVs (lower, %) under the assumption that both the ESW and MSS were stratified by only the survey period (alternative case)

Year/Block	1WW-M	1WW-L	1WM-H	1WM-M	1WM-L	1WE-H	1WE-M	1WE-L	1E-H	1E-M	1E-L	2-H	2-M
Latitude	25°N-35°	10°N-25°N	35°N-43°N	25°N-35°N	10°N-25°N	31°N-43°N	20°N-31°N	10°N-20°N	32°N-43°N	20°N-32°N	10°N-20°N	35°N-43°N	25°N-35°N
Longitude	130°B-137°30'E	130°B-137°30'E	137°30'E-145°E	137°30'E-145°E	137°30'E-145°E	145°E-165°E	145°E-165°E	145°E-165°E	165°E-180°	165°E-180°	165°E-180°	180°-155W	180°-155W
Area (nm^2)	188,870	385,344	81,960	233,552	384,389	692,020	663,632	692,984	309,600	581,514	519,738	289,250	778,545
1988						**1,726** 52.87			**676** 56.49			**1,167** 43.71	
1989						**2,433** 47.75	**1,277** 50.57		**2,913** 56.53	**977** 64.67		**225** 67.05	**1,145** 80.75
1990						**1,387** 76.30	**1,016** 55.64		**1,014** 44.85	**1,917** 51.06		**292** 59.60	**556** 113.39
1991			**0**										
1992	**0**			**0**		**2,687** 44.86							
1993		**2,688** 58.04			**999** 56.18	**4,451** 128.02	**4,350** 53.39	**3,785** 71.07	**9,158** 73.23	**4,791** 130.29	**418** 73.24		
1994													
1995											**0**		
1996								**4,375** 44.95					
Recent surveys (Abeam closing)													
1998						**3,851** 72.84	**1,721** 52.59	**169** 111.33					
1999									**3,254** 56.08	**9,078** 73.14	**295** 110.80		
2000			**0**	**0**	**235** 105.59								
2001	**0**	**369** 87.90											
2002												**973** 50.44	**1,030** 80.01
Recent surveys (Normal closing)													
1998						**747** 41.63	**3,205** 48.38	**475** 57.15					
1999									**3,194** 34.11	**5,699** 48.95	**258** 63.18		
2000			**0**	**0**	**502** 70.56								
2001	**0**	**0**											
2002												**939** 69.82	**1,430** 49.04

(c) Observed school counts and associated efforts (nm) used in a count-based model.

Year/Block	1WW-M	1WW-L	1WM-H	1WM-M	1WM-L	1WE-H	1WE-M	1WE-L	1E-H	1E-M	1E-L	2-H	2-M
Latitude	25°N-35°	10°N-25°N	35°N-43°N	25°N-35°N	10°N-25°N	31°N-43°N	20°N-31°N	10°N-20°N	32°N-43°N	20°N-32°N	10°N-20°N	35°N-43°N	25°N-35°N
Longitude	130°B-137°30'E	130°B-137°30'E	137°30'E-145°E	137°30'E-145°E	137°30'E-145°E	145°E-165°E	145°E-165°E	145°E-165°E	165°E-180°	165°E-180°	165°E-180°	180°-155W	180°-155W
Area (nm^2)	188,870	385,344	81,960	233,552	384,389	692,020	663,632	692,984	309,600	581,514	519,738	289,250	778,545
Past surveys (Normal closing)													
1988						**21** 4,948			**6** 1,616			**10** 1,457	
1989						**25** 4,181	**7** 2,139		**33** 2,062	**3** 1,050		**2** 1,513	**5** 1,999
1990						**13** 3,813	**2** 768		**9** 1,615	**9** 1,605		**1** 582	**2** 1,646
1991			**0** 3,157										
1992	**0** 2,233			**0** 2,243		**37** 5,601							
1993		**10** 843			**5** 1,131	**14** 1,279	**11** 987	**17** 1,830	**38** 755	**12** 856	**3** 2,194		
1994													
1995											**0** 2,261		
1996								**14** 1,304					
Recent surveys (Abeam closing)													
1998						**22** 1,431	**12** 1,676	**1** 1,487					
1999									**34** 1,171	**66** 1,531	**1** 637		
2000			**0** 230	**0** 122	**2** 1,185								
2001	**0** 111	**2** 756											
2002												**6** 646	**4** 1,095
Recent surveys (Normal closing)													
1998						**3** 1,006	**21** 1,574	**2** 1,056					
1999									**29** 1,018	**31** 1,145	**1** 729		
2000			**0** 81	**0** 94	**4** 1,109								
2001	**0** 153	**0** 1,125											
2002												**5** 558	**5** 986

Table 2. Specification of run sets for the estimation of the additional variance (AV) and CLA abundance in 1998-2002. Runs#1-8 given here correspond to those in Appendix 8 of SC59 Report ANNEX D, and Run#9 was added by following its suggestion. AV is estimated using data from Past period surveys (1988-1996) as well as Recent ones (1998-2002). Meanwhile, the point estimates of abundances in 1998-2002 are finally provided based only on data from Recent surveys, which was subject to IWC oversight, by making a correction of difference between the abeam and normal closing surveys. Therefore, the abundance estimates in 1998-2002 depend on the data from Past surveys in terms of their CVs. Remark that the correction factor was estimated twice separately in two estimation stages.

Run #	AV	Data included in the estimation of AV	Error model for AV	Covariates in the estimation of AV	R in the estimation of AV (Eq.1)	ESW / MSS	R* in the estimation of abundance in 1998-2002 (Eq.14)
1	Estimated	Past(NC) and Recent(AC&NC)	LN	Case1-4 (selected by AIC)	Estimated	By Sub-area	Estimated
2	Estimated	Past(NC) and Recent(AC&NC)	LN	Case1-4 (selected by AIC)	Estimated	Common	Estimated
3	Estimated	Past(NC) and Recent(AC&NC)	LN	Selected in Run#1	Fixed at 1	By Sub-area	Fixed at 1
4	N/A	-	-	-	-	(a) By Sub-area (b) Common	Estimated
5	Estimated	Past(NC) and Recent(AC&NC)	LN	Case5	Estimated	(a) By Sub-area (b) Common	Estimated
6	Estimated	Past(NC) and Recent(NC)	LN	Selected in Run#1	N/A	By Sub-area	Estimated
7	Estimated	Past(NC) and Recent(AC&NC)	LN	Selected in Run#1	Estimated	By Sub-area	Fixed at the estimate in AV model
8	Estimated	Past(NC)	LN	Case1	N/A	By Sub-area	Estimated
9	Estimated	Past(NC) and Recent(AC&NC)	NB	Case1-5 (selected by AIC)	Estimated	By Sub-area	Estimated

Characterization of Run sets

Run#1: Base scenario (based on the selected set of ESWs and mean school sizes)
Run#2: To know sensitivity to a different set of ESWs and mean school sizes)
Run#3: To know sensitivity to separation of survey modes
Run#4: To compare CVs in abundance estimates with/without consideration of additional variance
Run#5: For a generalized case of Case 4 (see below) to allow interactions between blocks and period
Run#6: To estimate the additional variance using data from only normal closing mode, where no correction factor is necessary
Run#7: To examine the impact of use of different estimates of the correction factor "R"
Run#8: To investigate the possibility to estimate additional variance based only on Past period surveys
Run#9: To assess the impact of ignoring the data with no sighting

Data included in the estimation of additional variance

Abeam Closing (AC)=closing starts after abeam of a school detected (conducted only in 1998-2002 surveys)
Normal Closing (NC)=closing starts soon after detection of a school (conducted in both of Past (1988-1996) and Recent (1998-2002) surveys)
Remark that the point estimates of abundances in Recent survey period for CLA purpose are estimated based only on the Recent surveys.

Error model for the estimation of additional variance

LN: the vector of logarithms of abundance estimates is assumed to be distributed to a multivariate normal distribution
NB: the numbers of detected schools given efforts, ESW and MSS are assumed to be independently distributed to negative distributions

Covariates in the estimation of additional variance

Case1: Block-effect
Case2: Block-effect and exponential trend
Case3: Block-effect and Period-effect
Case4: Block-effect and Period-Latitude interaction
Case5: Block-effect and Period-block interaction

ESW / mean school size (MSS)

Both the ESW and MSS were estimated by period. Sub-area was selected as a covariate for both the ESW and MSS in both the periods. Models with no covariates were secondarily selected in each period for each of ESW and MSS, and hence these were used in Run#2 as a sensitivity test.

Table 3: Results on the abundance estimates and their associated CVs under Run #1, #4(a) and #5(a).

Run#1 is the base cases, where the selected set of ESW and MSS was used in the underlying abundance estimates. "No additional variance" and "Case 5" correspond to Run #4(a) and #5(b), respectively. The estimation was done by two stages: the estimation of the additional CV, and then the estimation of abundance in 1998-2002. In the former stage, the additional CV was estimated using the data from Past and Recent surveys by the REML method. AIC was defined by the likelihood of the additional variance model. Meanwhile, the point estimates of abundances in 1998-2002 in the latter stage were estimated based only on the data from Recent surveys although their CVs depends on the values of additional CVs estimated in the former stage. Case 4 (Block-effect and Period-Latitude interaction) was selected as a covariate set for the additional variance model.

	No additional variance		**Case 1**		**Case 2**		**Case 3**		**Case 4**		**Case5**	
Parameter estimates and AIC in additional variance model												
			Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	**Estimate**	**CV(%)**	Estimate	CV(%)
R			0.953	24.85	1.004	26.17	0.891	26.17	**0.868**	**26.07**	0.892	26.55
Additional CV			0.627	31.44	0.638	31.94	0.641	31.47	**0.447**	**43.09**	0.588	43.48
AIC			21.58		22.84		22.75		**12.58** *		19.68	
Abundance in 1998-2002 survey period												
	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	**Estimate**	**CV(%)**	Estimate	CV(%)
R*	0.861	27.02	0.861	27.02	0.861	27.02	0.861	27.02	**0.861**	**27.02**	0.861	27.02
Block 1WW-L	348	88.08	348	121.39	348	122.40	348	122.72	**348**	**106.32**	348	117.90
Block 1WM-L	439	58.81	439	93.52	439	94.52	439	94.83	**439**	**78.40**	439	90.08
Block 1WE-H	1238	41.46	1238	79.52	1238	80.54	1238	80.86	**1238**	**63.71**	1238	75.97
Block 1WE-M	2525	38.57	2525	77.46	2525	78.49	2525	78.81	**2525**	**61.49**	2525	73.89
Block 1WE-L	407	53.61	407	89.06	407	90.06	407	90.38	**407**	**73.79**	407	85.60
Block 1E-H	3480	36.10	3480	75.79	3480	76.82	3480	77.14	**3480**	**59.67**	3480	72.20
Block 1E-M	7418	44.69	7418	81.92	7418	82.93	7418	83.25	**7418**	**66.28**	7418	78.40
Block 1E-L	315	56.65	315	91.64	315	92.64	315	92.96	**315**	**76.46**	315	88.19
Block 2-H	1813	48.41	1813	84.81	1813	85.82	1813	86.13	**1813**	**69.34**	1813	81.32
Block 2-M	2518	52.25	2518	87.93	2518	88.93	2518	89.25	**2518**	**72.61**	2518	84.46
Sub-area 1W	4957	28.03	4957	48.51	4957	49.08	4957	49.26	**4957**	**39.78**	4957	46.54
Sub-area 1E	11213	34.86	11213	60.91	11213	61.63	11213	61.86	**11213**	**49.83**	11213	58.41
Sub-area 2	4331	42.01	4331	65.61	4331	66.30	4331	66.51	**4331**	**55.28**	4331	63.26
Total	20501	25.49	20501	40.01	20501	40.43	20501	40.56	**20501**	**33.66**	20501	38.56

Correlation between abundance estimates

	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2
Sub-area 1W	0.289	0.195	0.095	0.072	0.093	0.070	0.093	0.070	0.142	0.105	0.103	0.078
Sub-area 1E	1	0.187	1	0.068	1	0.067	1	0.067	1	0.099	1	0.074
Sub-area 2		1		1		1		1		1		1

Table 4: Results on the abundance estimates and their associated CVs under Run #2, #4(b) and #5(b).

This table shows results under an alternative of Table 3 when another set of ESW and MSS were used. As in the Table 3, Case 4 (Block-effect and Period-Latitude interaction) was selected as a covariate set for the additional variance model.

	No additional variance		**Case 1**		**Case 2**		**Case 3**		**Case 4**		**Case5**	
Parameter estimates and AIC in additional variance model												
			Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	**Estimate**	**CV(%)**	Estimate	CV(%)
R			0.943	25.10	1.007	26.44	0.897	26.40	**0.881**	**26.29**	0.889	26.75
Additional CV			0.720	28.79	0.716	29.92	0.746	28.90	**0.533**	**35.87**	0.572	45.25
AIC			26.27		26.99		27.91		**17.90** *		18.45	
Abundance in 1998-2002 survey period												
	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	**Estimate**	**CV(%)**	Estimate	CV(%)
R*	0.858	27.23	0.858	27.23	0.858	27.23	0.858	27.23	**0.858**	**27.23**	0.858	27.23
Block 1WW-L	369	87.90	369	130.06	369	129.70	369	132.66	**369**	**112.97**	369	116.33
Block 1WM-L	466	58.57	466	101.94	466	101.60	466	104.45	**466**	**85.12**	466	88.48
Block 1WE-H	1314	41.13	1314	88.04	1314	87.70	1314	90.57	**1314**	**70.80**	1314	74.30
Block 1WE-M	2679	38.18	2679	86.00	2679	85.66	2679	88.54	**2679**	**68.65**	2679	72.18
Block 1WE-L	432	53.36	432	97.50	432	97.16	432	100.01	**432**	**80.60**	432	84.00
Block 1E-H	3561	35.29	3561	84.11	3561	83.76	3561	86.65	**3561**	**66.63**	3561	70.20
Block 1E-M	7592	44.03	7592	90.15	7592	89.80	7592	92.67	**7592**	**73.01**	7592	76.48
Block 1E-L	322	56.18	322	99.88	322	99.54	322	102.39	**322**	**83.03**	322	86.40
Block 2-H	1071	41.07	1071	88.00	1071	87.65	1071	90.53	**1071**	**70.76**	1071	74.26
Block 2-M	1488	45.36	1488	91.15	1488	90.80	1488	93.67	**1488**	**74.05**	1488	77.50
Sub-area 1W	5260	27.51	5260	53.10	5260	52.91	5260	54.53	**5260**	**43.45**	5260	45.39
Sub-area 1E	11475	34.01	11475	66.57	11475	66.33	11475	68.38	**11475**	**54.33**	11475	56.80
Sub-area 2	2558	33.22	2558	65.40	2558	65.16	2558	67.19	**2558**	**53.32**	2558	55.76
Total	19293	25.80	19293	45.09	19293	44.94	19293	46.21	**19293**	**37.62**	19293	39.11

Correlation between abundance estimates

	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2
Sub-area 1W	0.359	0.308	0.095	0.081	0.096	0.082	0.090	0.077	0.142	0.122	0.130	0.111
Sub-area 1E	1	0.289	1	0.075	1	0.076	1	0.071	1	0.113	1	0.103
Sub-area 2		1		1		1		1		1		1

Table 5: Results on the abundance estimates and their associated CVs under Run #3, #6, #7 and #8.

Run#3 was conducted to know sensitivity to separation of survey modes. Run#6 aimed at estimating the additional variance using data from only normal closing mode, where no correction factor is necessary. Also, Run#7 was to examine the impact of use of different estimates of the correction factor in the final abundance estimates. In Run#8, only abundance estimates from Past survey were employed, and therefore Case 1, where only block-effects were considered, was assumed as the covariates. The other three runs were based on Case 4, which was selected through Run#1 and Run#5. Remark that, in Run#3 and Run#7, the estimation uncertainty of the correction factor was not taken into account in CVs in the final abundance estimates, and therefore their CVs were not comparable with those in other runs.

	Run#3		**Run#6**		**Run#7**		**Run#8**	
Parameter estimates and AIC in additional variance model								
	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)
R	1	-	-	-	0.868	22.64	-	-
Additional CV	0.450	42.75	0.503	40.22	0.447	43.10	0.588	43.48
Abundance in 1998-2002 survey period								
	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)
R*	1	-	0.861	27.02	0.868	**0**	0.861	27.02
Block 1WW-L	348	106.55	348	110.72	348	106.32	348	117.90
Block 1WM-L	439	75.55	439	82.89	439	75.31	439	90.08
Block 1WE-H	1238	59.76	1238	68.48	1238	59.49	1238	75.97
Block 1WE-M	2525	59.45	2525	66.32	2525	59.18	2525	73.89
Block 1WE-L	407	69.93	407	78.34	407	69.68	407	85.60
Block 1E-H	3480	55.73	3480	64.56	3480	55.46	3480	72.20
Block 1E-M	7418	63.07	7418	70.98	7418	62.82	7418	78.40
Block 1E-L	315	73.05	315	80.98	315	72.80	315	88.19
Block 2-H	1813	68.61	1813	73.98	1813	68.36	1813	81.32
Block 2-M	2518	69.19	2518	77.18	2518	68.94	2518	84.46
Sub-area 1W	4957	36.44	4957	42.40	4957	36.21	4957	46.54
Sub-area 1E	11213	45.81	11213	53.14	11213	45.56	11213	58.40
Sub-area 2	4331	52.66	4331	58.35	4331	52.64	4331	63.25
Total	20501	28.74	20501	35.54	20501	28.70	20501	38.56

Correlation between abundance estimates

	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2
Sub-area 1W	0	0	0.125	0.093	0	0	0.104	0.078
Sub-area 1E	1	0	1	0.088	1	0	1	0.074
Sub-area 2		1		1		1		1

Table 6: Results on the abundance estimates and their associated CVs under Run #9

In this run, instead of assuming the log-normality of abundance estimates, negative binomial distributions were supposed for the block-wise counts of detected schools. This aimed at assessing the impact of ignoring the data with no sighting, which were eliminated from the data in runs other than Run#9. In this run, however, "0" data in the three blocks (1WW-M, 1WM-H, 1WM-M) were still eliminated because no sighting were made in those blocks through the whole two periods, and hence no information on the additional variance was drawn from those data. The method for providing the abundance in 1998-2002 was same as other runs. This run also selected Case 4 as the covariate for the additional variance model. Since the estimate of the additional CV in Case 4 was larger than that in Run#1 (base case), CVs in the final estimates were also greater than those in the base case.

	Case 1		**Case 2**		**Case 3**		**Case 4**		**Case5**	
Parameter estimates and AIC in additional variance model										
	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	**Estimate**	**CV(%)**	Estimate	CV(%)
R	0.929	23.65	0.979	26.05	0.871	25.15	**0.875**	**26.75**	0.873	25.19
Additional CV	0.778	34.12	0.789	34.84	0.784	35.17	**0.535**	**50.24**	0.719	41.82
AIC	257.0		258.3		258.2		**248.2 ***		255.0	
Abundance in 1998-2002 survey period										
	Estimate	CV(%)	Estimate	CV(%)	Estimate	CV(%)	**Estimate**	**CV(%)**	Estimate	CV(%)
R*	0.861	27.02	0.861	27.02	0.861	27.02	**0.861**	**27.02**	0.861	27.02
Block 1WW-L	348	136.09	348	137.16	348	136.67	**348**	**113.30**	348	130.16
Block 1WM-L	439	107.77	439	108.80	439	108.33	**439**	**85.49**	439	102.07
Block 1WE-H	1238	93.92	1238	94.95	1238	94.48	**1238**	**71.21**	1238	88.20
Block 1WE-M	2525	91.92	2525	92.95	2525	92.48	**2525**	**69.08**	2525	86.18
Block 1WE-L	407	103.32	407	104.35	407	103.88	**407**	**80.97**	407	97.63
Block 1E-H	3480	90.29	3480	91.32	3480	90.85	**3480**	**67.35**	3480	84.55
Block 1E-M	7418	96.27	7418	97.29	7418	96.83	**7418**	**73.68**	7418	90.56
Block 1E-L	315	105.90	315	106.92	315	106.46	**315**	**83.59**	315	100.20
Block 2-H	1813	99.11	1813	100.14	1813	99.67	**1813**	**76.64**	1813	93.41
Block 2-M	2518	102.20	2518	103.23	2518	102.76	**2518**	**79.82**	2518	96.50
Sub-area 1W	4957	56.60	4957	57.19	4957	56.92	**4957**	**43.90**	4957	53.38
Sub-area 1E	11213	71.16	11213	71.89	11213	71.56	**11213**	**55.06**	11213	67.07
Sub-area 2	4331	75.39	4331	76.10	4331	75.77	**4331**	**60.12**	4331	71.47
Total	20501	46.01	20501	46.44	20501	46.25	**20501**	**36.63**	20501	43.61

Correlation between abundance estimates

	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2	SA 1E	SA 2
Sub-area 1W	0.095	0.072	0.093	0.070	0.093	0.070	0.142	0.105	0.103	0.078
Sub-area 1E	1	0.068	1	0.067	1	0.067	1	0.099	1	0.074
Sub-area 2		1		1		1		1		1

Table 7. Correlation matrices among the final estimates of the abundance in blocks and the correction factor. Correlations were made through the common ESWs and MSSs in the same sub-areas as well as the common correction factor (*R**). Remark that the abundance in 1WW-L and those in blocks in sub-areas 1E and II were not correlated because the schools were detected in 1WW-L only by the abeam closing mode, and therefore the abundance was solely estimated by the underlying abundance estimate by the abeam closing mode (therefore no correction factor was used for this estimate).

Case 4 in Run#1

	1WW-L	1WM-L	1WE-H	1WE-M	1WE-L	1E-H	1E-M	1E-L	2-H	2-M	R*
1WW-L	1										
1WM-L	0.007	1									
1WE-H	0.008	0.052	1								
1WE-M	0.008	0.042	0.054	1							
1WE-L	0.007	0.047	0.061	0.049	1						
1E-H	0.000	0.043	0.057	0.043	0.052	1					
1E-M	0.000	0.037	0.049	0.037	0.045	0.067	1				
1E-L	0.000	0.035	0.046	0.035	0.042	0.063	0.055	1			
2-H	0.000	0.019	0.025	0.019	0.023	0.026	0.023	0.021	1		
2-M	0.000	0.037	0.048	0.036	0.044	0.049	0.043	0.040	0.109	1	
R*	0.000	-0.170	-0.223	-0.169	-0.203	-0.229	-0.199	-0.188	-0.105	-0.194	1

Case 4 in Run#9

	1WW-L	1WM-L	1WE-H	1WE-M	1WE-L	1E-H	1E-M	1E-L	2-H	2-M	R*
1WW-L	1										
1WM-L	0.006	1									
1WE-H	0.007	0.046	1								
1WE-M	0.007	0.037	0.047	1							
1WE-L	0.006	0.042	0.054	0.043	1						
1E-H	0.000	0.038	0.049	0.037	0.045	1					
1E-M	0.000	0.033	0.043	0.032	0.039	0.058	1				
1E-L	0.000	0.031	0.041	0.031	0.038	0.055	0.049	1			
2-H	0.000	0.017	0.022	0.017	0.020	0.023	0.020	0.019	1		
2-M	0.000	0.032	0.042	0.032	0.039	0.043	0.037	0.036	0.096	1	
R*	0.000	-0.161	-0.208	-0.158	-0.191	-0.213	-0.186	-0.177	-0.098	-0.183	1

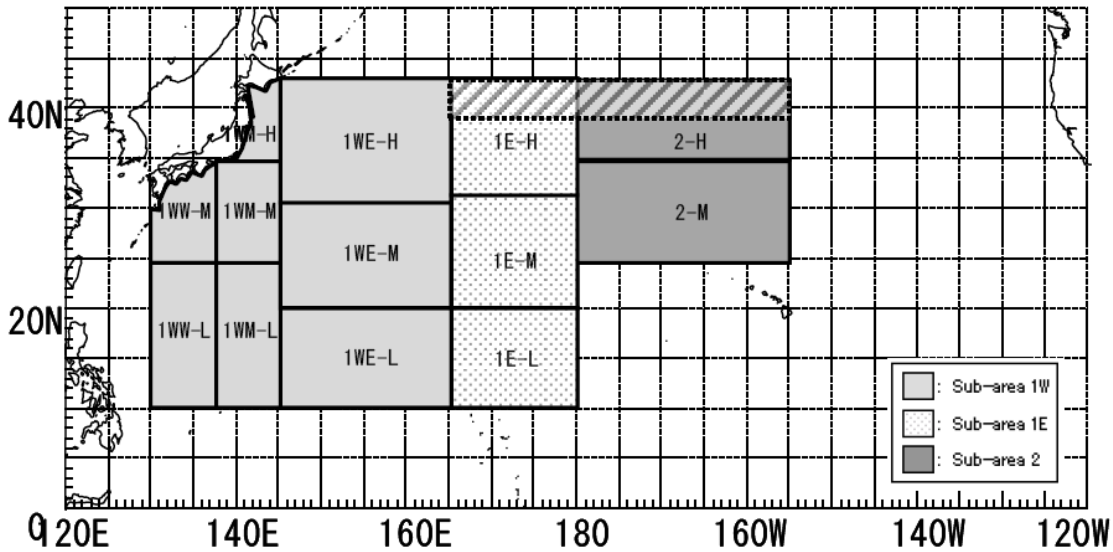
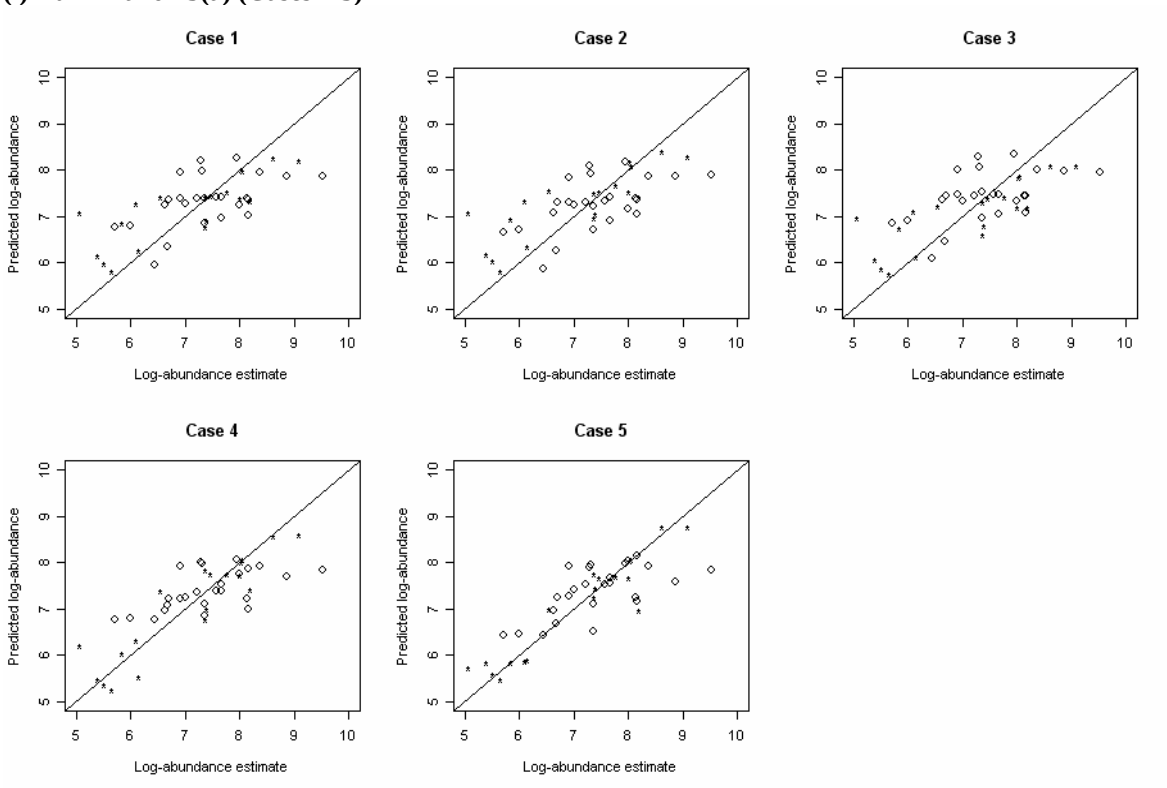


Figure 1. A map of the sub-areas and blocks used for the abundance estimation. “H”, “M” and “L” mean high, middle and low latitudes. The northern parts (shaded) in the two blocks, 1E-H and 2-H, were excluded from the estimation of abuncances, which means any detections and efforts in those parts were not included in the analyses and the abundance estimates in those blocks were calculated for the southern parts of 1E-H and 2-H. A more detailed explanation is given in Shimada et al. (2008)

(i) Run #1 and #5(a) (Cases 1-5)



(ii) Run #9 (Cases 1-5)

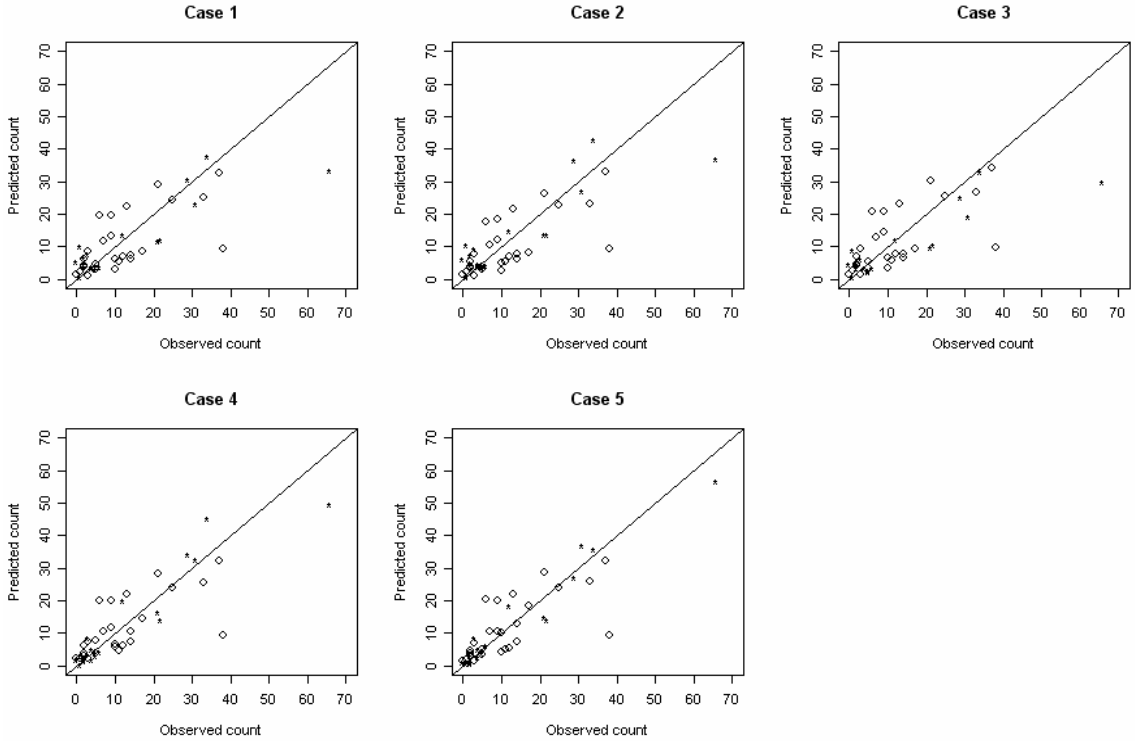
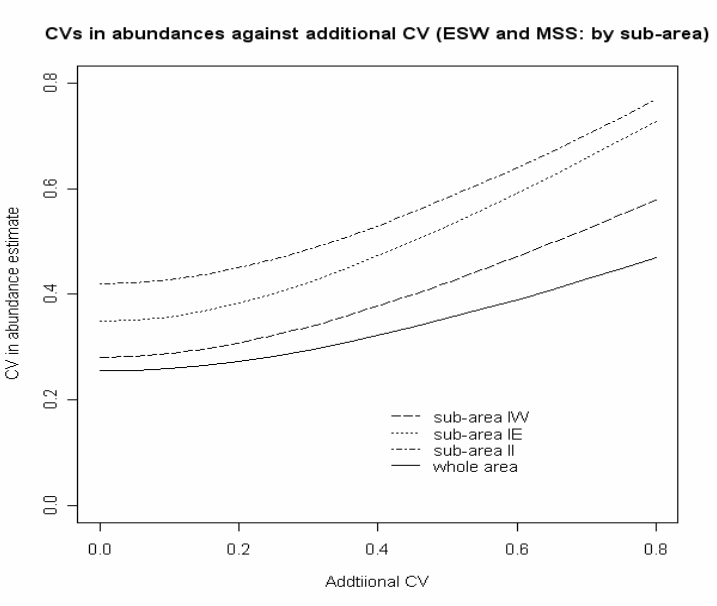


Figure 2. Observed and predicted values for Run #1 and Run #5(a) (Cases 1-5) and Run#9 (Cases 1-5); circles for the past surveys and asterisks for the recent surveys.



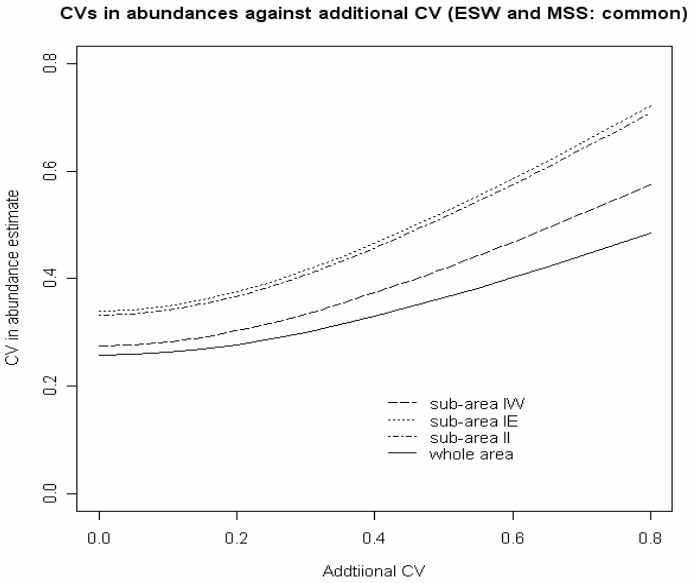


Figure 3. Plots of CVs in abundance estimates against the value of the additional CV.

Appendix: Replies to "SUGGESTIONS FOR ADDITIONAL ANALYSES/DOCUMENTATION RELATED TO THE ESTIMATION OF ABUNDANCE FOR THE WESTERN NORTH PACIFIC BRYDE'S WHALES" in Appendix 8 in Annex D in SC59 report.

1) Passing versus closing mode

a. The estimates of closing/passing mode density estimates of R~0.92-0.97 in the integrated analyses are difficult to reconcile with the differences in the overall sighting rates between passing and closing mode (i.e. n/l in passing mode is 1.51 larger than n/l in closing mode), especially given that common esw and school size estimates are used in the integrated analyses. This requires examination to provide some explanation. (Extracted from Annex D)

> Under the data set revised in Shimada et al. (2008), the differences in the overall sighting rates between passing and closing mode was 1.308. The reciprocal of this value, 0.765, was slightly different from the estimate R^* in Case 4 in Run#1 (around 0.86). However, the value R^* was estimated with taking the difference in sampling CVs among blocks into account, and therefore it was essentially assessed as weighted mean. So, these two values were comparable but slightly different.

b. Provide an estimate of R based on the data from the 1998-2002 surveys alone and compare it with those from "integrated" analysis. (Extracted from Annex D)

> In Case 4 in Run#1, where the model captured the behavior of the data well, the values of R in the integrated model and R^* estimated based only on the recent surveys were almost same (R=0.868 and R^*=0.861). This was because in this case the information on these correction factors were drawn from the difference in the underlying abundance estimates between two survey modes (abeam and normal closing) in "the recent surveys". By way of comparison , for example in Case 1, where abundance levels were assumed to be same between the two survey periods, this was not the case (R=0.953 and R^*=0.861) because the abundance estimates in the past surveys contributed to the estimation of fixed block-effects. Anyway, in out selected cases, no clear differences were observed between the two correction factors.

2) The random effects approach to estimating abundance / alternative analyses

a. Table 1 (in Appendix 8 in Annex D) lists the current set of models, and an alternative model (case 5) that generalizes case 4 to allow for interactions between block and period. It also include a series of alternative additional analyses to allow an examination of: (a) whether the inclusion of the 1988-96 data impact the point estimates of abundance for 1998-2002, (b) the impact of fixing rather than estimating R, (c) the impact of basing the analyses on closing mode data only (because there are no data for passing mode for the years 1988-96), and (d) allowing esw and mean school size to differ among areas (i.e. in the context of SC/59/PF3 Small Areas). (Extracted from Annex D)

b. There are blocks with some effort and no sightings for the 1988-96 period. The impact of ignoring the data for these blocks, particularly in terms of the impact of the extent of additional variance, should be more fully examined and documented in the final report. Analyses which involve delta-log-normal models or an error model (such as the negative binomial) which allows for zero observations should provide information on this. (Extracted from Annex D)

The nine of run sets shown in Table 2 in our paper reflected these points.

c. The process of calculating the variance-covariance matrix for the abundance estimates should be more fully documented. In particular, the procedure to take account of the covariance resulting from using a single pooled esw and mean school size estimate should be more fully documented. (Extracted from Annex D)

d. Show the variance-covariance matrix (for the selected case) to be able to assess the correlation between R and the other parameters. (Extracted from Annex D)

The formulas for assessing the variance-covariance matrix were provided in the text (e.g., (2), (3), (17) and (19)). Also, the actual matrices were given as the correlation matrices for selected cases in Table7.