Consideration of measurement errors for the Norwegian common minke whale (*Balaenoptera acutorostrata acutorostrata*) surveys

Hiroko K. Solvang[1], Hans J. Skaug[1,2] and Nils Øien[1]

Contact e-mail: hirokos@hi.no

ABSTRACT

A discrete measurement error (ME) model for radial distance and angle to detected objects in line transect surveys is considered. This approach directly quantifies the effect of ME on the estimated effective strip half-width. The method is applied to data from experiment and sighting data collected over the period 2008–2013 in North Atlantic both under the assumption of multiplicative and additive ME. The results indicate that the abundance estimates obtained by taking ME into account are consistently larger than the abundance estimates without ME correction. Furthermore, some insights into the effect of ME are provided by a theoretical analysis.

KEYWORDS: ABUNDANCE ESTIMATE; ATLANTIC OCEAN; MINKE WHALE; SURVEY-VESSEL; NORTH ATLANTIC

INTRODUCTION

Distance-based abundance estimation methods such as line-transect sighting surveys *inter alia* require robust and unbiassed estimates of the position of the animal relative to the trackline of the vessel (e.g. Buckland *et al.*, 1993).

The Norwegian vessel line-transect surveys for common minke whale use a double platform design (observers search for and record whale sightings from two independent observer platforms at different heights on the vessel) in which the sighting angle is measured on an angle board and the radial distance is estimated by eye for not only the initial sighting but also for each subsequent surfacing of a whale (Skaug *et al.*, 2004). Following guidelines established by the International Whaling Commission's Scientific Committee for whale sightings surveys (e.g. IWC, 2012), distance and angle experiments are routinely conducted as part of the surveys with the aim of estimating bias and variability in measurement error (ME) so that this can be incorporated into the analyses to obtain abundance estimates and their associated uncertainty (Bøthun *et al.*, 2008). A simulation-based correction method has previously been applied to the abundance estimates (Skaug *et al.*, 2004) that also accounted for other factors, such as errors in duplicate identification. However, the isolated effect of distance and angle ME was not explicitly quantified. In a more general setting, Marques (2004) derived the probability density of the observed distances for a continuously distributed error. Borchers *et al.* (2010) considered an extension of that method and provided a likelihood that accounts for ME.

In the present paper, the probability density of radial distance is implemented directly using a multiplicative ME model approach following Marques (2004) and Borchers *et al.* (2010). For sighting angles, an additive error model is also considered since this may be more appropriate when the error tends not to increase with the original angle. For computational simplicity, discrete ME models are used. The approach here takes into account the double-platform design and surface tracking protocol that has been developed for the Norwegian common

[1] Institute of Marine Research, PO Box 1870 Nordnes, N-5817 Bergen, Norway.
[2] University of Bergen, N-5020, Bergen, Norway.

minke whale abundance surveys. The ME correction is applied both to the initial sightings and to the subsequent sightings and the effects are compared. Finally, theoretical results (asymptotic expansions) are provided on the effect of ME on radial distance density. These results are also discussed in terms of providing a qualitative insight into ME.

MATERIALS AND METHODS

Data

In this study we used two data sources: distance and angle experiments, and sightings data.

Distance and angle experiments

The distance and angle experiments were conducted midway and at the end of the survey while training was performed incidentally throughout the survey. The experiment was conducted by using two radar-reflective buoys as target objects under sighting conditions representative of the actual survey (Beaufort sea state 2–3) but with a visible horizon. The buoys were dropped into the sea at distances of 1,000 to 4,000 meters from the vessel. The vessel then moved towards the buoys at a speed between 6 and 10 knots following courses determined by the cruise leader.

For the experiment, one observer was present at a time on each platform. The upper platform (Platform 1) was typically a barrel on the mast whilst the lower platform (Platform 2) was on the wheelhouse roof. On a signal from the cruise leader, the observer provided an 'instant' estimate of the angle (degrees) from a mounted angle board and distance (meters) to the specified buoy by naked eye. At the same time, the true values were recorded by the ship's radar operator. All measurements were recorded in the same way as to a whale sighting during the survey. This procedure was repeated several times as the vessel approaches the buoys. The cruise leader switched between the buoys at random and the vessel also changed the course several times during the approach towards the buoys to reduce possible serial correlation. About 10 readings were obtained per observer (and of course a set of true measurements from the radar). Once one round was completed, the vessel moves away from the buoys and new observers were tested following the same procedure. The analyses covered the 2008–2013 combined survey estimate for Northeast Atlantic minke whales (Solvang *et al.*, 2015), and data from four participating vessels denoted as ERO, THO, JHJ and HGU were used. Data from 2013 which used different methodology were excluded. The data include 524 observations for distance from Platform 1 and 541 observations from Platform 2, while there were 532 and 550 angle observations from Platform 1 and Platform 2, respectively. Variable numbers of data points in trials result from observers not always being able to collect both angle and distance measurements successfully.

Sighting data from 2008-2013

The sighting data was collected by annual surveys (except for 2012) between 2008–2013, covering the Northeast Atlantic from the North Sea (southern boundary 52°N) to the ice edge, and from the Greenland Sea in the west to the Barents Sea in the east.

Whales were searched for by naked eye from two platforms each manned with two observers in line with the experimental methodology. The survey and sightings protocols were: Primary searching speed was intended to be 10 knots and the surveys were conducted in passing mode. When searching, one observer in the team scanned the port 45° sector from the transect line while the other scanned the starboard 45° sector. Sightings made outside these sectors and all initial sightings before abeam have been used in the analyses. Acceptable conditions for primary searching were defined as meteorological visibility greater than 1km and Beaufort sea state of 4 or less.

For each sighting, radial distance as estimated by eye, angle from the transect line as read from an angle board, school size and swimming direction were reported. Specific tracking procedures were followed, following the whale where possible reporting positional data (radial distance, angle) of all subsequent surfacings until the whale passed, or was assumed to have passed, behind abeam. The units of observation are the tracks of observed surfacings.

Data on weather conditions, Beaufort sea state, meteorological sightability and glare were recorded regularly on an hourly basis or when conditions changed notably.

Over the survey period 2008–2013 a total of 779 sightings of groups (sum of platforms A and B) were made during primary search effort.

Correction for measurement error

In Skaug *et al.* (2004), the key parameters for abundance calculation were estimated based on the likelihood for observations made by both platforms. In the current paper we incorporate a model for measurement error into the likelihood. This model is estimated using the experiment data. Then we estimate the effective strip half-width for the sighting data under different assumptions about the measurement error model.

For initial sightings made by only one of the platforms, say Platform 1, any subsequent surfacing detected by Platform 1 will set up a Bernoulli trial for Platform 2 (still unaware of the whale's presence) with dichotomous outcome (seen/not seen). Combined across all detected whales, initially seen by either Platform 1 or 2, these trials give rise to a Bernoulli likelihood which is combined with the likelihood for the initial sightings (Skaug *et al.*, 2004).

Multiplicative measurement error

Consider first a single platform and denote by r_t and r_o the true and recorded (observed) radial distance, respectively. For each individual observation, errors of the form observed = true × error are referred to as being multiplicative and are calculated as $c^{(r)} = r_t/r_o$, so that $c^{(r)} = 1$ means no error ($r_t = r_o$). Similarly, for angle measurements, the error is calculated as $c^{(\theta)} = \theta_t/\theta_o$ where θ_t is the true sighting angle and θ_o is the observed sighting angle). Although it is the joint probability density of (r_t, θ_o) that is needed for the likelihood (Skaug *et al.*, 2004), for simplicity we will start with the marginal density of r_o.

Marques (2004) introduced the multiplicative ME model, although with a slightly different notation than used here. Both Marques (2004) and Borchers *et al.* (2010) assumed a continuous distribution for the multiplicative ME. The density of observed radial distances then becomes

$$f_o(r_o) = \int f_t(cr_o) f_c(c) c \cdot dc \ , \tag{1}$$

where f_o, f_t and f_c are the probability density of r_o, r_t and $c^{(r)}$, respectively. For computational simplicity, we assume a discrete distribution for $c^{(r)}$. A discretised version of eq. (1) is

$$f_o(r_o) = \sum_{i=1}^{k} f_t(c_i r_o) p(c_i) c_i \ , \tag{2}$$

where $c_1 < c_2 < \cdots < c_k$ are appropriately chosen support points of the probability mass function $p(c)$. The corresponding probabilities must satisfy $p(c_1) + \cdots + p(c_k) = 1$. Eq. (2) is hence a discrete mixture distribution model.

The joint probability of r_o and θ_o is required to construct the likelihood function for the initial sightings (Skaug *et al.* 2004) . The version of (2) which applies to (r_o,θ_o) is

$$f_o(r_o,\theta_o) = \sum_{i=1}^{k} f_i\left(c_i^{(r)} r_o, c_i^{(\theta)} \theta_o\right) \times p^{(r)}(c_i^{(r)}) \times c_i^{(r)} \times p^{(\theta)}(c_i^{(\theta)}) \times c_i^{(\theta)} \ , \tag{3}$$

where $f_t(r_o,\theta_o)$ is given as the product of (2) and (5) in Skaug *et al.* (2004) and see Appendix.

Additive measurement error

When exploring the effect of additive ME model for sighting angle, the error is defined as $c^{(\theta)} = \theta_t - \theta_o$, and eq. (3) becomes

$$f_o(r_o,\theta_o) = \sum_{i=1}^{k} f_i\left(c_i^{(r)} r_o, c_i^{(\theta)} + \theta_o\right) \times c_i^{(r)} p^{(r)}(c_i^{(r)}) \times p^{(\theta)}(c_i^{(\theta)}) \ . \tag{4}$$

Likelihood of the model

To construct the likelihood function for the initial sightings (Skaug *et al.*, 2004), the joint probability density is given by (3) in the case (case I) of multiplicative ME for both distance and angle and given by (4) in the case (case II) of multiplicative ME for distance and additive ME for angle.

Bernoulli trials only apply to a single platform. If the whale previously has been sighted by Platform 1, say, subsequent sightings made by 1 will generate Bernoulli trials for Platform 2 (until the whale is detected). The dichotomous outcome variable (success/failure) is denoted by d, and $f(d \mid r,\theta)$ is used to emphasise that the probability of d depends on the relative position (r,θ) where the trials takes place. The position (r,θ), which is determined by Platform 1, is also affected by ME. The Bernoulli probabilities for the cases I and II, conditionally on the observed position, are

$$f(d \mid r_o,\theta_o) = \sum_{i=1}^{k} f_i(d \mid c_i^{(r)} r_o, c_i^{(\theta)}\theta_o) \times p^{(r)}(c_i^{(r)}) \times p^{(\theta)}(c_i^{(\theta)}) \,, \tag{5}$$

and

$$f(d \mid r_o,\theta_o) = \sum_{i=1}^{k} f_t\left(d \mid c_i^{(r)} r_o, c_i^{(\theta)} + \theta_o\right) \times p^{(r)}(c_i^{(r)}) \times p^{(\theta)}(c_i^{(\theta)}) \,, \tag{6}$$

respectively, where $f_t(r_t,\theta_t)$ is as in eq (3) given as the product of (2) and (5) described in Skaug *et al* (2004) (see Appendix, A).

Setting for discrete measurement error

In this study, a discretised error model with only three levels for each of the terms $c^{(r)} = r_t/r_o$ and $c^{(\theta)} = \theta_t/\theta_o$ or $c^{(\theta)} = \theta_t - \theta_o$ is assumed. That is, $k = 3$ and $p(c_1) + p(c_2) + p(c_3) = 1$, for each of $c^{(r)}$ and $c^{(\theta)}$, where the center point $c_2 = 1$ corresponds to no ME. Values for the parameters c_1, c_3, $p(c_1)$, $p(c_2)$, $p(c_3)$ were obtained from the data from experiment as follows. In the case of $c^{(r)}$ the empirical distribution of $c^{(r)}$ was split into three classes according to the 5%, 90%, and 5% percentiles (other percentile values, e.g. 10%, 80%, 10% ,were used as well). Then, c_1 was taken to be the average value within the leftmost class, and similarly c_3 was taken as the average within the rightmost class. The corresponding class probabilities were $p(c_1) = 0.1$, $p(c_2) = 0.8$ and $p(c_3) = 0.1$. We next perform a theoretical analysis of the effect of discrete ME.

Asymptotic expansions

If it is assumed that $k = 3$, $c_1 = 1 - \varepsilon$, and $c_3 = 1 + \varepsilon$, the density (2) becomes

$$f_o(r_o) = (1-\varepsilon)p_1 f_t((1-\varepsilon)r_o) + p_2 f_t(r_o) + (1+\varepsilon)p_3 f_t((1+\varepsilon)r_o) \,, \tag{7}$$

where p_1, p_2, and p_3 correspond to $p(c_1)$, $p(c_2)$ and $p(c_3)$. Here, $\varepsilon > 0$ is some small quantity, so we are investigating the limiting case of no measurement error ($\varepsilon = 0$). The objective is to try to provide some insights into the effect of measurement error on the density f_o, which is useful for understanding the implied effect on the effective strip half-width.

It may be shown mathematically that

$$\begin{aligned} f_o(r_o) = f_t(r_0) &+ (p_3 - p_1)\left\{f_t(r_0) + r_0 f_t^{'}(r_o)\right\}\cdot\varepsilon \\ &+ 2(p_1 + p_3)\left\{r_0 f_t^{'}(r_0) + \frac{1}{2} r_0^2 f_t^{''}(r_0)\right\}\cdot\varepsilon^2 + o(\varepsilon^2) \,, \end{aligned} \tag{8}$$

where $f_t^{'}$ and $f_t^{''}$ denote the first and second derivative of the function f_t, and $o(\varepsilon^2)$ represents terms that can be ignored when ε is small. To ease the interpretation of eq. (8) an exponential probability distribution is assumed, given by

$$f_t(r_o) = \frac{1}{\mu} e^{-\frac{r_o}{\mu}} \,, \tag{9}$$

where μ is a parameter (mean) of the exponential distribution. Inserting eq. (9) into (8) gives

$$f_0(r_0) \approx \frac{1}{\mu} e^{-\frac{r_0}{\mu}} + \frac{(p_3 - p_1)}{\mu} e^{-\frac{r_0}{\mu}} \left(1 - \frac{r_0}{\mu}\right) \cdot \varepsilon + \frac{1}{2} \cdot \frac{(p_1 + p_3) r_0}{\mu^2} e^{-\frac{r_0}{\mu}} \left(-2 + \frac{r_0}{\mu}\right) \cdot \varepsilon^2 + o(\varepsilon^2) \,. \quad (10)$$

For a situation with $p_1 > p_3$ (i.e. a positively biased distance measurement) for sufficiently small $\varepsilon, f_0(r_0) < f_t(r_0)$ for $r_0 < \mu$ and $f_0(r_0) > f_t(r_0)$ for $r_0 > \mu$, i.e. the probability mass is moved to larger distances. Similarly, when $p_1 < p_3$ the probability mass is moved to smaller distances. When $p_1 = p_3$, i.e. there is no systematic component to the measurement error, the first order (ε) term of the right-hand side of eq. (10) is zero, and the second order (ε^2) term needs to be considered. This term is negative when $r_0 < 2\mu$ which means that the probability mass is moved to larger values as an effect of random (unbiased) measurement error. The parameter μ was estimated from the experimental data by maximum likelihood using the function '*fitdistr*' in library *MASS* of R (*https://cran.r-project.org*).

Measurement error and the estimation of the effective strip half-width

To obtain insights into the effect of measurement error on effective strip half-widths (*eshw* – a key parameter in obtaining abundance estimates, the multiplicative or additive ME models were fitted to experiments conducted on all vessels, based on an estimate of the effective strip half-width given by equation (1) and the detection function given by equation (3) in Skaug *et al.* (2004) (see Appendix, B). Four different settings are considered and *eshw* was calculated for one platform:

(a) ME for case I (see 'Likelihood of the model' in Material and Methods) is accounted for only for the initial sighting;

(b) ME for case I is accounted for both in initial sighting and Bernoulli trials;

(c) ME for case II is accounted for only in initial sighting; and

(d) ME for case II is accounted for both in initial sighting and Bernoulli trials.

To check the sensitivity to the choice of *k* in eq. (2), we also conducted the same procedure for *k* = 5, where the corresponding class probabilities were $p(c_1) = p(c_2) = p(c_3) = p(c_4) = p(c_5) = 0.2$.

RESULTS AND DISCUSSION

Configuration for experimental measurement errors

Using experimental data, histograms for the multiplicative ME of radial distance are shown in Fig. 1, and histograms for the multiplicative and additive measurement errors for angle are shown in Figs 2 and 3, respectively. The empirical mean values and the coefficient of variation (CV) have also been added in each histogram.

Despite the observed individual differences between vessels, the overall result is a tendency for underestimating distances for $c = r_t/r_o > 1$, i.e. the observed distance is less than the true distance. Fig. 4 shows histograms of true and observed distances on Platform 1 pooled across vessels, as well as a scatter plot between them. Based on the information in Fig. 4, underestimation of radial distances seems to take place up to around 600m and above 2,000m while in the interval 600m to 2,000m there is a tendency to overestimate radial distance. In the analyses of whale sightings data (Solvang *et al.*, 2015), radial distances were truncated at 2,000m and the experimental radial distance data reflect those observed during the survey quite well (Solvang *et al.*, 2015; Fig. 2).

As expected, the distributions for multiplicative and additive angle errors were mostly centered around 1 as seen in Fig. 2 and around 0 as seen in Fig. 3. Deviations from this are hard to interpret as the error will depend on absolute values (less sensitivity to higher values). In addition, when travelling towards the buoy in the same direction, the angle will increase by delay in reading, which will have a greater effect the nearer the vessel is to the buoy and cause the ratio to be less than 1. The additive measurement error represents the angle precision

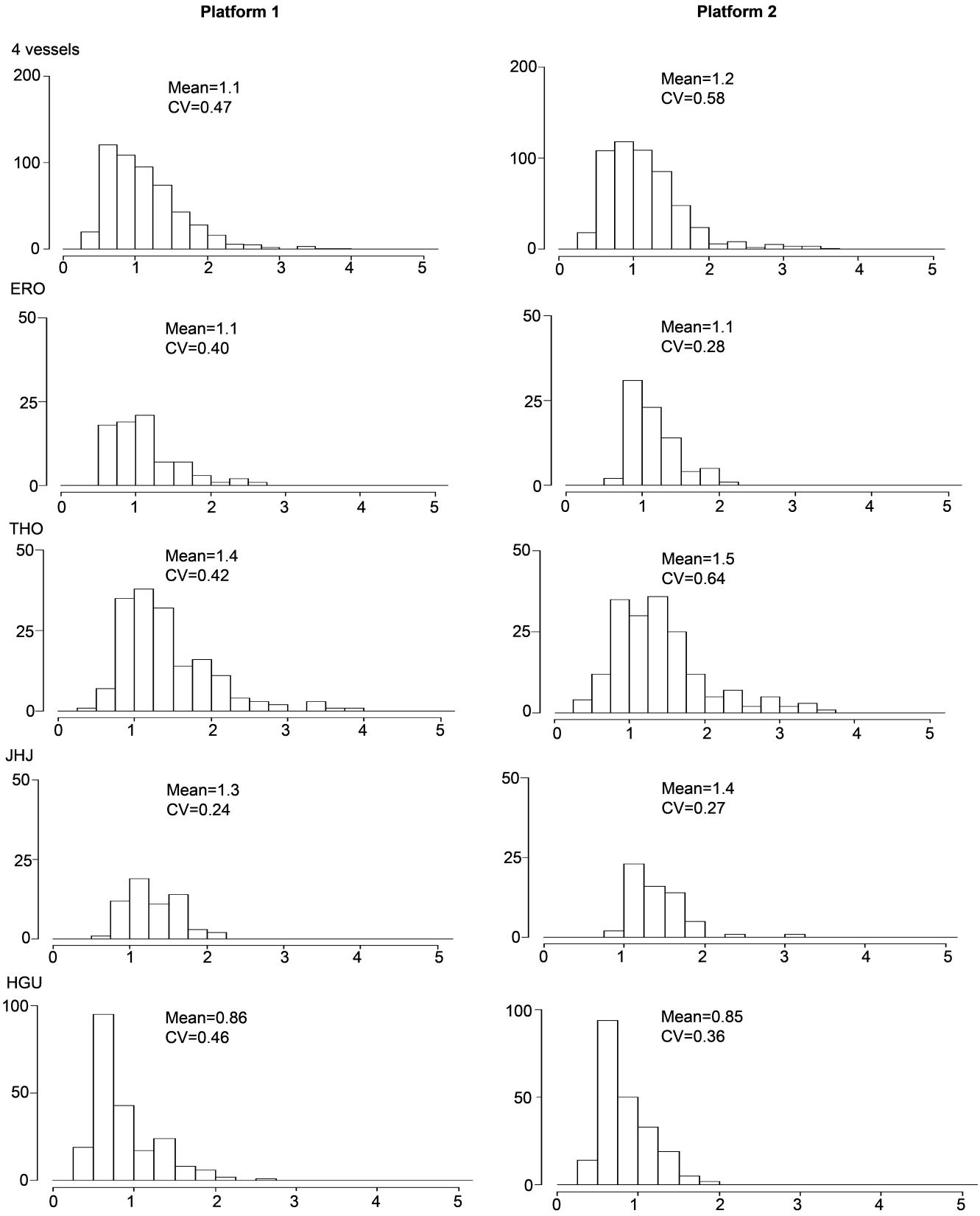

Fig. 1. Distribution of multiplicative measurement error ($c^{(r)} = r_t/r_o$) for radial distance by vessel and platform. The top row shows data aggregated over the four vessels.

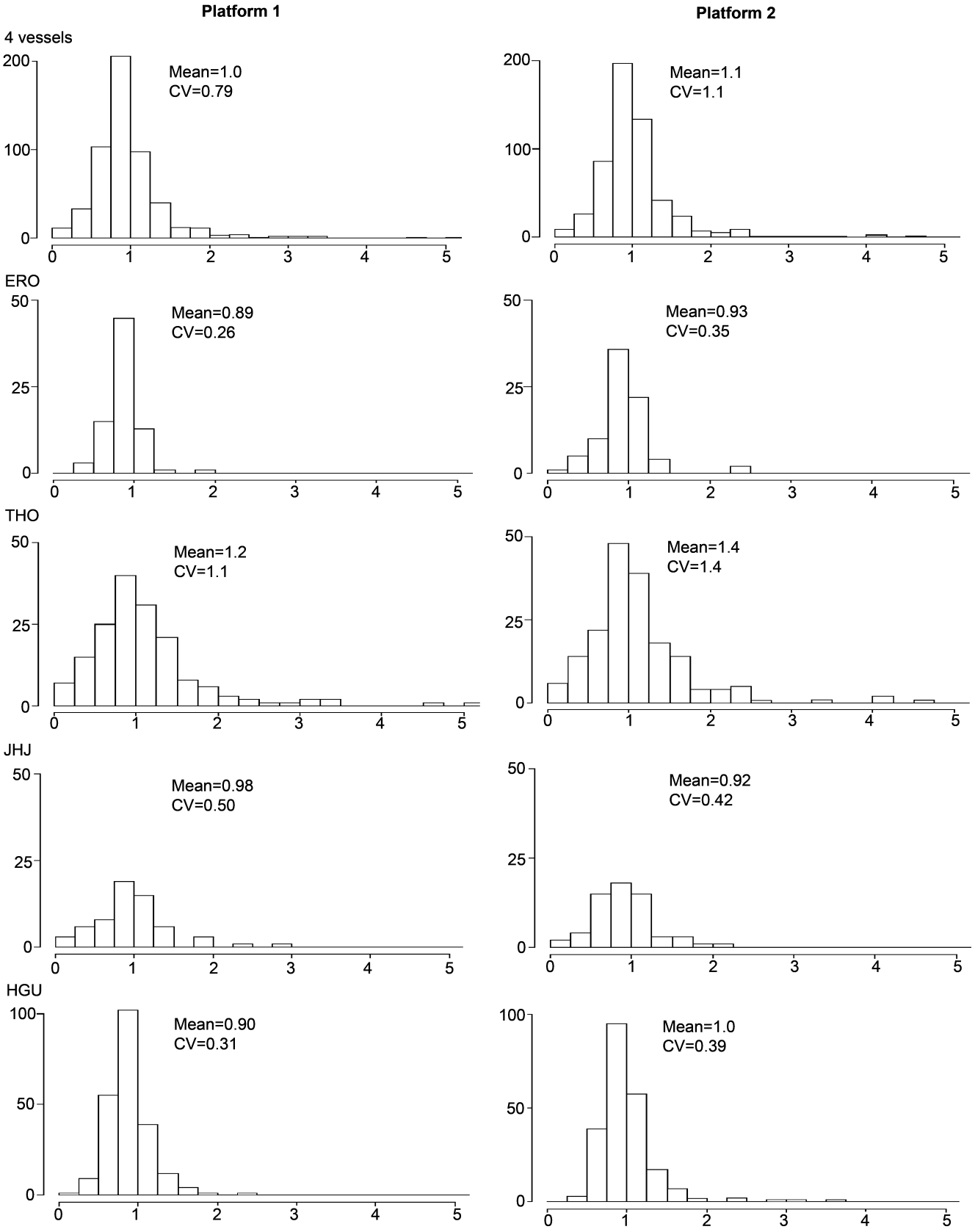

Fig. 2. Distribution of multiplicative measurement error ($c^{(\theta)} = \theta_t/\theta_o$) for sighting angle by vessel and platform. The top row shows data aggregated over the four vessels.

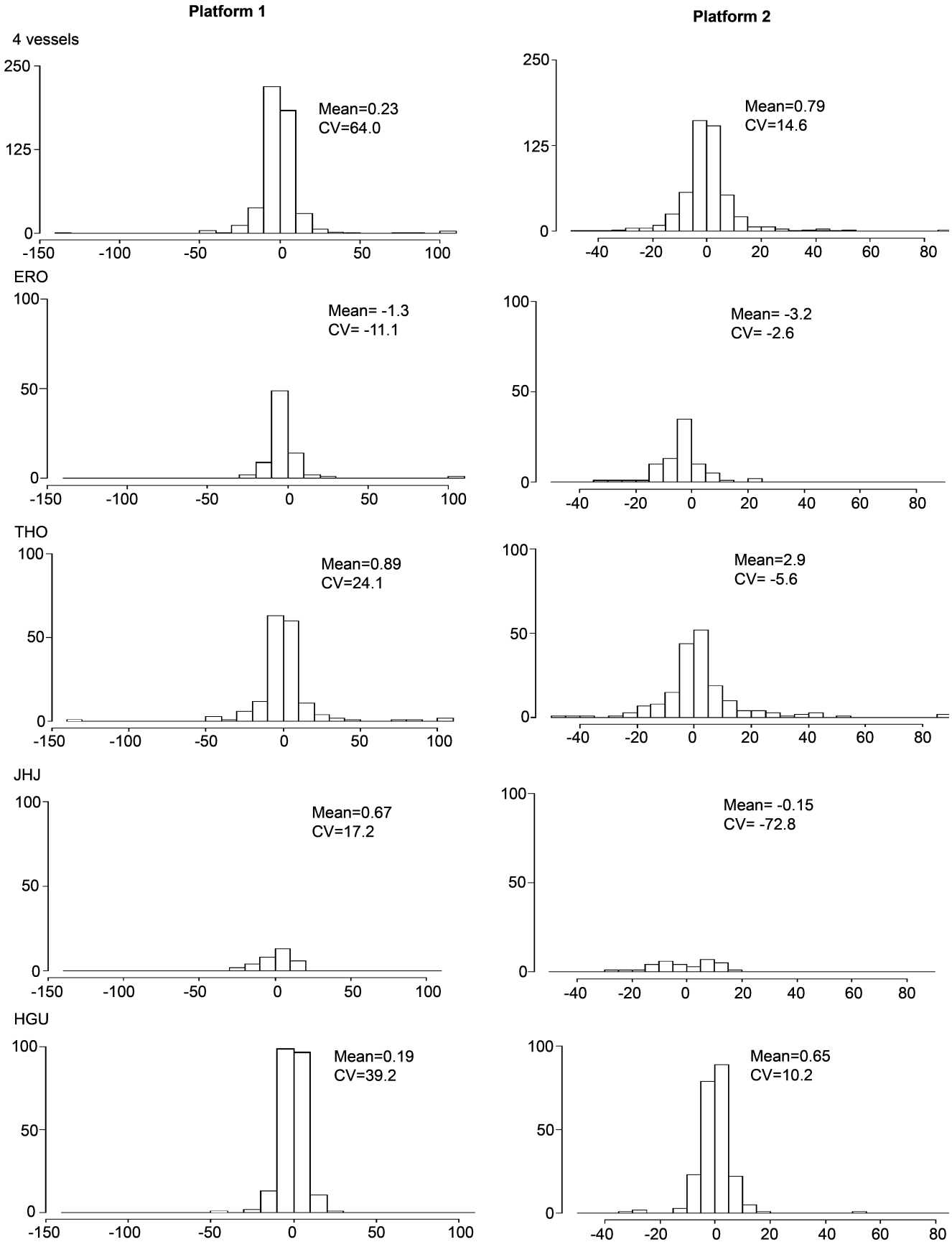

Fig. 3. Distribution of additive measurement error ($c^{(\theta)} = \theta_t - \theta_o$) for sighting angle by vessel and platform. The top row shows data aggregated over the four vessels.

better as the histograms would be expected to be symmetrical around zero. This is supported by three panels shown in Fig. 5, which summarise histograms of true and observed angles and scatter plots across them on Platform 1 of all vessels. However, small discrepancies may be introduced into the measurements due to the fact that the two platforms may not be completely aligned to the reference point. Since the mean values for additive error cases become less than the standard deviation, the CV will be larger than in the multiplicative cases and thus the two CVs cannot be directly compared.

Fig. 4 shows that radial distance is underestimated at distances r > 1,000m. Sighting angle, on the other hand, appears unbiased for all angles (Fig 5).

Applied measurement error models and asymptotic expansions

Using the sighting data from 2008-2013, *eshw* for the four settings were calculated for investigating into the effects of ME. The results are summarised in Tables 1a–d. The tables include expected values and CV for distance

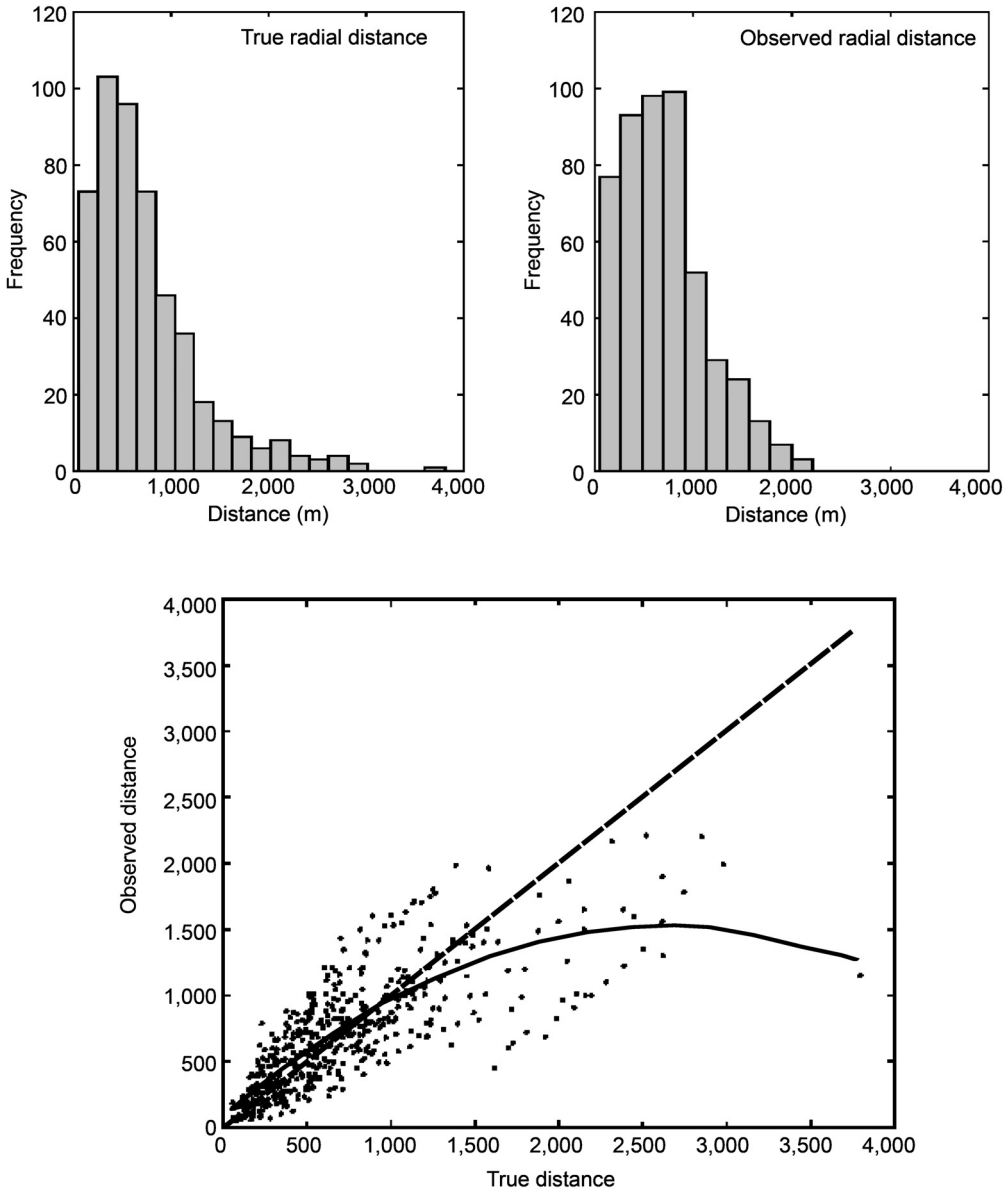

Fig. 4. Histograms of true radial distance r_t (upper left) and observed radial distance r_o (upper right) on platform 1 pooled across vessels. The lower panel shows a scatter plot between true observed distances, with the dashed line representing points with no measurement error ($c^{(r)} = 1$) and the solid line indicating loess smoothing plot to the scatter plot.

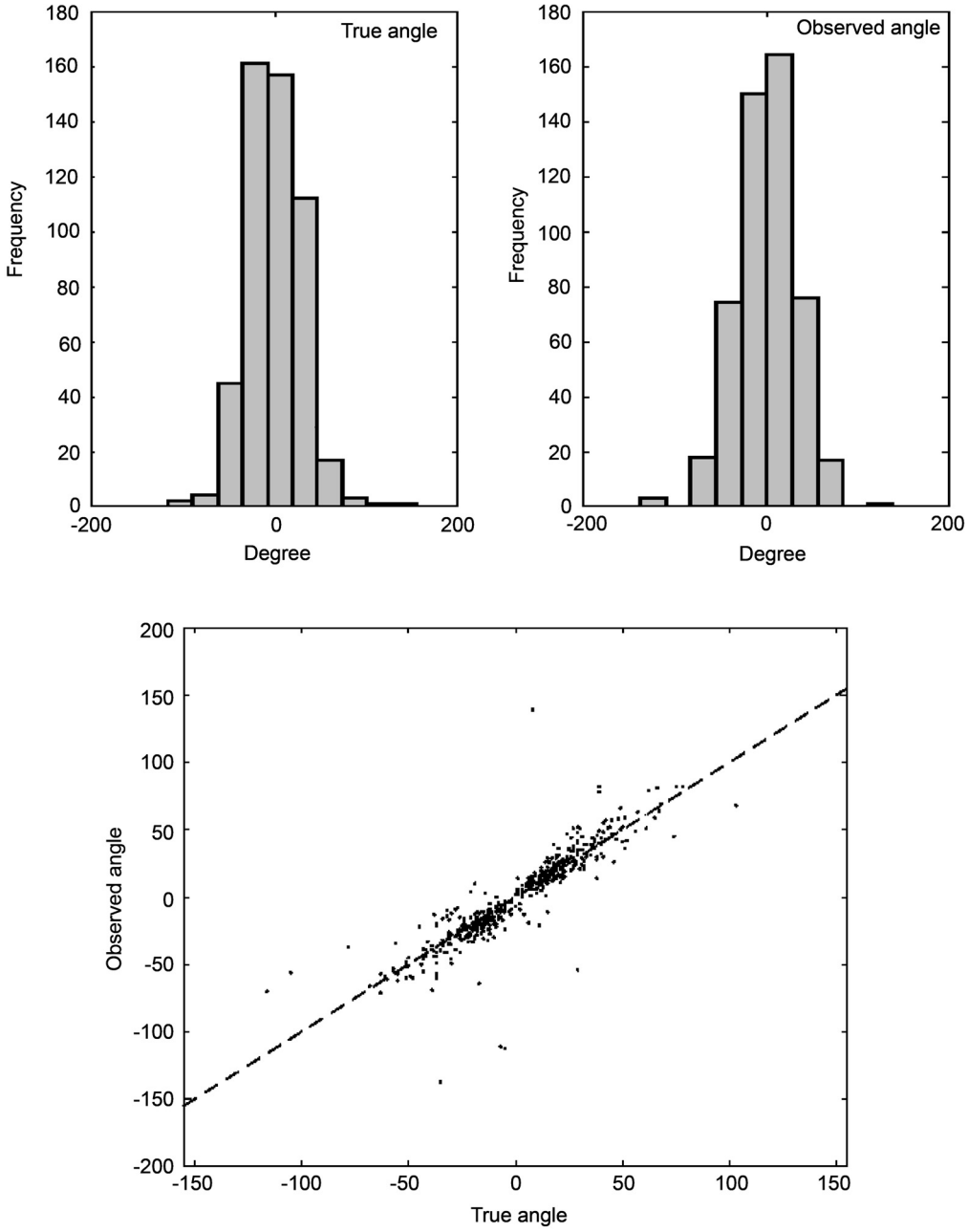

Fig. 5. Histograms of true angle θ_t (upper left) and observed radial angle θ_o (upper right) on platform 1 pooled across all vessels. The lower panel shows a scatter plot between true observed angle, with the dashed line representing points with no measurement error ($c^{(\theta)} = 1$).

and angle for each platform separately and average for both platforms (*eshw* was calculated for Platform 1). To check the sensitivity to the choice of *k* in eq. (2), the same procedure was conducted for *k* = 5 and the results are shown in the additional panels under each table.

Compared to the *eshw* assuming no ME (ranging from around 280m to 290m depending on setting) eshw with 10–40% measurement error was always less than without any measurement error. Furthermore, this remained the case if measurement error was assumed to occur only for either distance or angle (assuming 20% measurement error). Thus, the tendency for decreased *eshw* if measurement error is taken into account was consistent for all cases. A consistent tendency was observed for $c^{(r)} = r_o/r_t$. This can be seen in the two panels in Fig. 6, which illustrate the estimated density function to perpendicular distances according to settings b. and d. – results were similar for settings a. and c. although those are not shown here. The density function in the case of no ME (solid line) shows the longest tail i.e. leads to the largest *eshw*. A similar tendency is found when

Table 1

Effective strip half-widths (eshw) and the standard deviation in parentheses and for different levels of measurement error. Also shown is the average error (Mean c) and coefficient of variation of the error (CV c) for distance (r) and angle (ϑ). Upper panel, k = 3; lower panel, k = 5.

(a) Multiplicative error model both for distance and angle applied to initial likelihood only

					Distance (r)		Angle (ϑ)	
ME	p_1	p_2	p_3	eshw (std)	Mean c	CVc	Mean c	CVc
No ME	0%	100%	0%	280.3(15.3)	1.1	0.0	1.0	0.0
Observation for	10%	80%	10%	237.7(13.9)	1.1	0.37	1.0	0.48
platform 1	20	60	20	216.8(13.2)	1.1	0.40	1.0	0.44
	30	40	30	194.6(11.4)	1.1	0.40	1.0	0.39
	40	20	40	197.2(11.4)	1.1	0.40	1.0	0.38
Observation for	10%	80%	10%	253.4(14.8)	1.2	0.41	1.1	0.54
platform 2	20	60	20	235.9(14.3)	1.2	0.42	1.1	0.46
	30	40	30	242.1(14.7)	1.2	0.41	1.1	0.41
	40	20	40	242.7(14.2)	1.2	0.40	1.1	0.37
Average of platforms	10%	80%	10%	244.7(14.3)	1.2	0.39	1.1	0.51
1 and 2	20	60	20	227.6(13.8)	1.2	0.41	1.1	0.45
	30	40	30	238.3(14.3)	1.2	0.41	1.1	0.40
	40	20	40	212.4(12.3)	1.2	0.40	1.1	0.38
		20% ME only distance		240.1(14.2)	1.2	0.41	–	–
		20% ME only angle		266.0(14.9)	–	–	1.1	0.45

							Distance (r)		Angle (ϑ)	
	p_1	p_2	p_3	p_3	p_3	eshw	Mean c	CV c	Mean c	CV c
Platform 1						196.3(11.8)	1.1	0.43	1.0	0.44
Platform 2	20%	20%	20%	20%	20%	233.4(14.3)	1.2	0.44	1.1	0.47
Average of Platforms 1 and 2						212.4(12.3)	1.2	0.43	1.1	0.46

(b) Multiplicative error model for both distance and angle applied to both likelihood components

					Distance (r)		Angle (ϑ)	
ME	p_1	p_2	p_3	eshw (std)	Mean c	CVc	Mean c	CVc
No ME	0%	100%	0%	289.0(15.4)	1.1	0.0	1.0	0.0
Observation for	10%	80%	10%	240.0(14.0)	1.1	0.37	1.0	0.48
platform 1	20	60	20	213.6(13.2)	1.1	0.40	1.0	0.44
	30	40	30	188.8(11.2)	1.1	0.40	1.0	0.39
	40	20	40	192.4(11.1)	1.1	0.40	1.0	0.38
Observation for	10%	80%	10%	264.9(15.4)	1.2	0.41	1.1	0.54
platform 2	20	60	20	243.7(14.9)	1.2	0.42	1.1	0.46
	30	40	30	252.7(14.9)	1.2	0.41	1.1	0.41
	40	20	40	251.9(14.4)	1.2	0.40	1.1	0.37
Average of platforms	10%	80%	10%	251.1(14.6)	1.2	0.39	1.1	0.51
1 and 2	20	60	20	229.8(14.1)	1.2	0.41	1.1	0.45
	30	40	30	212.2(12.5)	1.2	0.41	1.1	0.40
	40	20	40	215.0(12.2)	1.2	0.40	1.1	0.38
		20% ME only distance		248.0(14.5)	1.2	0.41	–	–
		20% ME only angle		269.3(15.2)	–	–	1.1	0.45

							Distance (r)		Angle (ϑ)	
	p_1	p_2	p_3	p_3	p_3	eshw	Mean c	CV c	Mean c	CV c
Platform 1						191.0(11.6)	1.1	0.43	1.0	0.44
Platform 2	20%	20%	20%	20%	20%	242.2(14.6)	1.2	0.44	1.1	0.47
Average of Platforms 1 and 2						208.0(12.5)	1.2	0.43	1.1	0.46

(c) Multiplicative error model for distance and additive model for angle applied to initial likelihood only

ME	p_1	p_2	p_3	eshw (std)	Distance (r)		Angle (ϑ)	
					Mean c	CVc	Mean c	CVc
No ME	0%	100%	0%	280.4(15.3)	1.1	0.0	1.0	0.0
Observation for	10%	80%	10%	246.7(14.4)	1.1	0.37	0.8	13.3
platform 1	20	60	20	229.9(13.8)	1.1	0.40	1.2	8.0
	30	40	30	209.4(12.1)	1.1	0.40	1.3	7.0
	40	20	40	212.7(12.0)	1.1	0.40	0.8	11.6
Observation for	10%	80%	10%	248.1(14.5)	1.2	0.41	1.3	7.3
platform 2	20	60	20	239.3(14.2)	1.2	0.42	1.0	8.8
	30	40	30	231.2(13.5)	1.2	0.41	1.4	5.7
	40	20	40	228.1(12.9)	1.2	0.40	1.6	4.7
Average of platforms	10%	80%	10%	223.3(13.1)	1.2	0.39	1.0	9.6
1 and 2	20	60	20	235.8(14.0)	1.2	0.41	1.6	8.6
	30	40	30	223.3(13.1)	1.2	0.41	1.4	6.3
	40	20	40	224.2(12.7)	1.2	0.40	1.1	7.3
		20% ME only distance		240.3(14.2)	1.2	0.41	–	–
		20% ME only angle		275.3(15.1)	–	–	1.6	8.6

	p_1	p_2	p_3	p_3	p_3	eshw	Distance (r)		Angle (ϑ)	
							Mean c	CV c	Mean c	CV c
Platform 1						207.8(12.2)	1.1	0.43	1.0	9.4
Platform 2	20%	20%	20%	20%	20%	223.7(13.1)	1.2	0.44	1.3	7.0
Average of Platforms 1 and 2						218.1(12.8)	1.2	0.43	1.2	8.1

(d) Multiplicative error model for distance and additive model for angle applied to both likelihood components

ME	p_1	p_2	p_3	eshw (std)	Distance (r)		Angle (ϑ)	
					Mean c	CVc	Mean c	CVc
No ME	0%	100%	0%	289.8(15.5)	1.1	0.0	1.0	0.0
Observation for	10%	80%	10%	256.3(14.7)	1.1	0.37	0.8	13.3
platform 1	20	60	20	237.4(14.1)	1.1	0.40	1.2	8.0
	30	40	30	212.8(12.2)	1.1	0.40	1.3	7.0
	40	20	40	215.5(12.1)	1.1	0.40	0.8	11.6
Observation for	10%	80%	10%	265.1(15.1)	1.2	0.41	1.3	7.3
platform 2	20	60	20	249.7(14.5)	1.2	0.42	1.0	8.8
	30	40	30	241.8(13.6)	1.2	0.41	1.4	5.7
	40	20	40	237.7(13.1)	1.2	0.40	1.6	4.7
Average of platforms	10%	80%	10%	259.8(14.8)	1.2	0.39	1.0	9.6
1 and 2	20	60	20	244.6(14.3)	1.2	0.41	1.6	8.6
	30	40	30	230.5(13.1)	1.2	0.41	1.4	6.3
	40	20	40	232.5	1.2	0.40	1.1	7.3
		20% ME only distance		248.7(14.5)	1.2	0.41	–	–
		20% ME only angle		285.2(15.3)	–	–	1.6	8.6

	p_1	p_2	p_3	p_3	p_3	eshw	Distance (r)		Angle (ϑ)	
							Mean c	CV c	Mean c	CV c
Platform 1						212.1(12.3)	1.1	0.43	1.0	9.4
Platform 2	20%	20%	20%	20%	20%	232.5(13.3)	1.2	0.44	1.3	7.0
Average of Platforms 1 and 2						225.0(12.9)	1.2	0.43	1.2	8.1

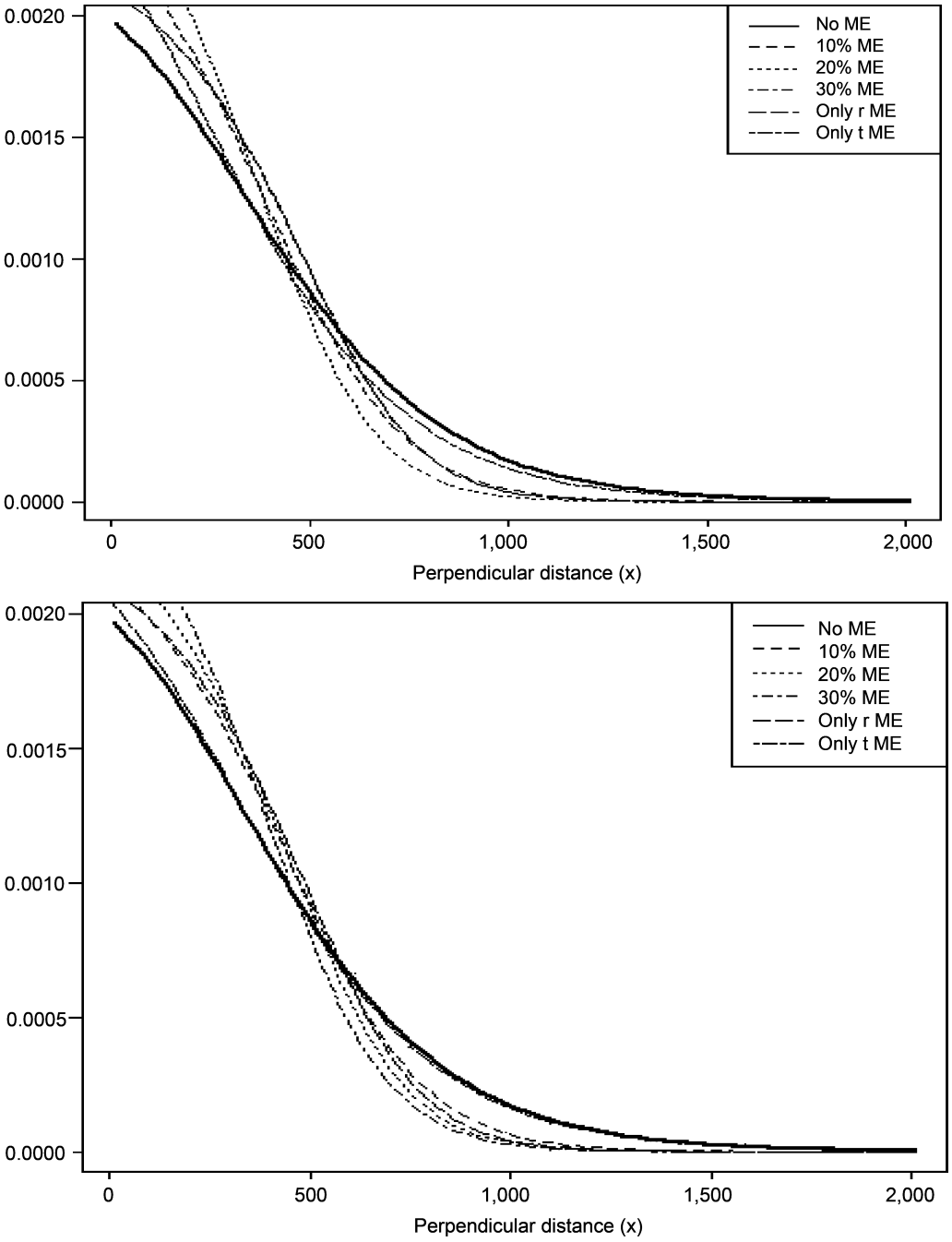

Fig. 6. Estimated perpendicular distance density for different ME levels for both multiplicative (upper) and additive (lower) ME model for sighting angle. The y-axis indicates the probability density of perpendicular distance $x = r \sin \theta$.

applying multiplicative ME to only distance. However, when additive ME is applied (only angle), the lines are much more similar to no ME than the case where multiplicative ME is applied to angle. This is in accord with the results in Tables 1a–d showing that the *eshw* for 20% multiplicative ME only angle is always less than the *eshw* for 20% additive ME only angle. It is caused by the difference between $\theta_t = \theta_o + 0.2$ and $\theta_t = 0.2\theta_o$. We assume simple assumption for $p_1 = p_3$, while Figs 1 and 2 indicate right tailed distribution that we assume $p_1 \neq p_3$. Nevertheless, the *eshw* would be less as long as $p_2 < 1$.

Abundance estimates are inversely proportional to the *eshw*. Since estimated *eshw* incorporating measurement error is less than if it is assumed there has been no measurement error, then abundance estimates

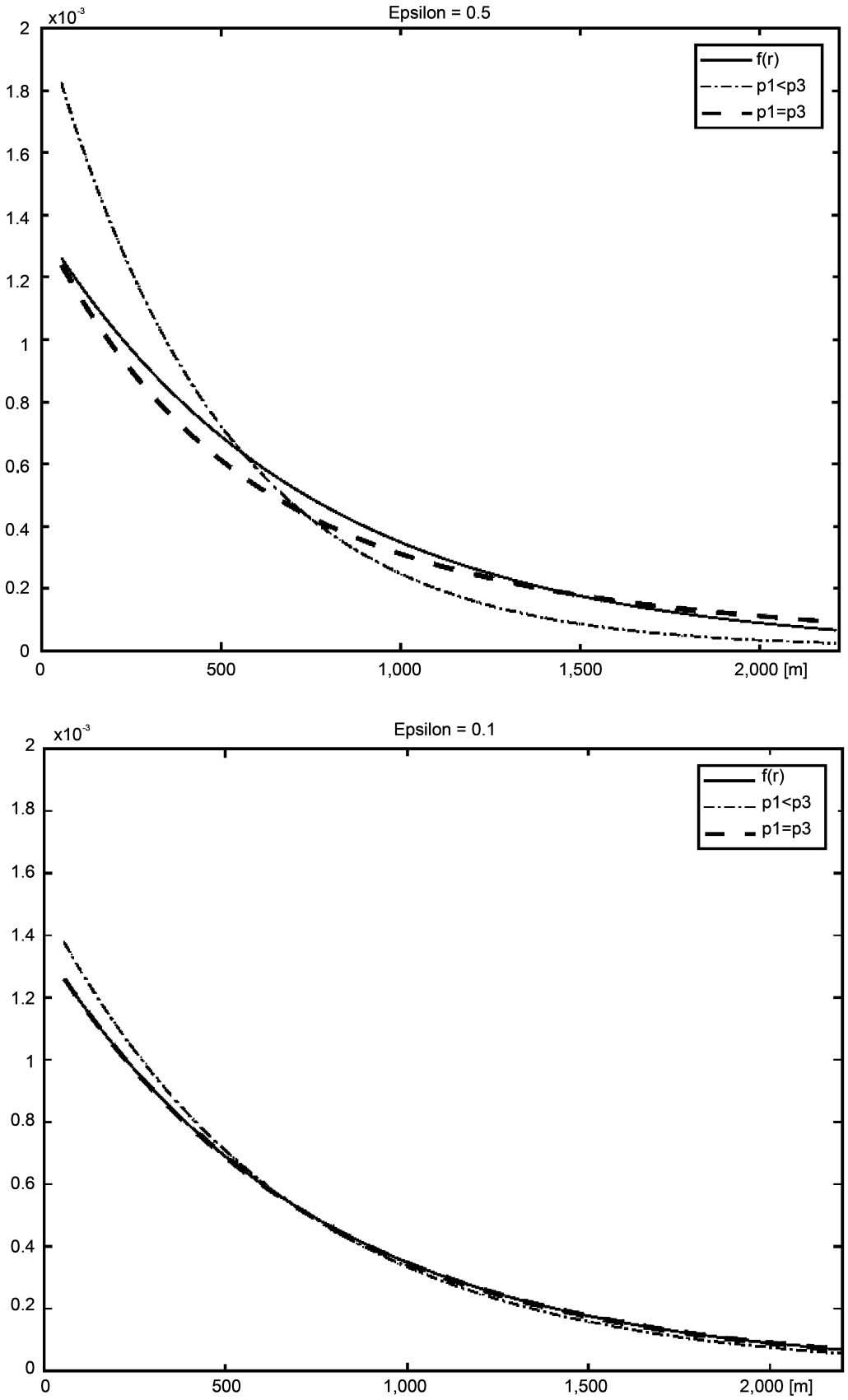

Fig. 7. Radial distance density $f_o(r)$ based on eqn. (11) with $\mu = 1.36 \times 10^{-3}$ for symmetric ($p_1 = p_3$) (and asymmetric ($p_1 < p_3$) distribution of c. Two different levels of ME in eqn. (10) is used: $\varepsilon = 0.5$ (upper) and $\varepsilon = 0.1$ (lower).

that are not corrected for measurement error are always lower than those for which the error is incorporated and thus are conservative estimates. This tendency is consistent with the experiments for the survey period 2002–2007 (Bøthun *et al.*, 2008).

Finally, we consider the asymptotically expansion of the distribution function for the observations with measurement error using true data for the radial distance and the observation from platform 1 in the experimental data (Fig. 4). The configuration of the histograms for true distance are not so different from one for the observed data. First, we identify the parameter μ to calculate $f_t(r_o)$ using the true data according to eq. (19) in Skaug *et al.* (2004). The estimated μ was 1.36×10^{-3} by the maximum likelihood fitting for the true data. As we see in Fig. 7, the condition $p_3 > p_1$ seems natural. This is because the histograms of true and observed distances have heavy right-hand tails. In Fig. 7 we illustrated the approximated $f_t(r_o)$, given by eq. (10), for $p_3 > p_1$ or $p_3 = p_1$ in the cases of $\varepsilon = 0.5$ and $\varepsilon = 0.1$. The approximated $f_t(r_o)$ for $\varepsilon = 0.1$ converges closer to the distance density rather than the case for $\varepsilon = 0.5$. This consideration suggests that the difference between the approximated density and the true density becomes extremely small if ε (measurement error) goes to zero.

CONCLUSION

We investigated the effect of measurement error on the estimated effective strip half-widths, a key parameter in obtaining abundance estimates, for the 2008–13 Norwegian minke whale survey estimate using a discrete ME model. For all models considered we found that the abundance estimates increased when correcting for ME. This increase is directly caused by a decrease in eshw. It should be kept in mind that the eshw is affected both by the distribution of perpendicular distances and the detection probability on the track line, commonly referred to as g(0). These two factors interact during the estimation process in a non-transparent way. Systematic bias in measured distances can inflate or deflate the apparent eshw in either direction, depending on the direction of the bias. Hence, our conclusion of a deflated abundance estimate as a consequence of ME is specific to the Norwegian minke whale surveys. Our mathematical demonstration that unbiased/random multiplicative ME shifts the perpendicular/radial distance distribution towards larger distances is, however, generally valid under the set of stated assumptions. Other factors which should be investigated further, are consideration of environmental factors like variation in Beaufort sea state and visibility of a permanently surfaced 'buoy' versus a real whale that could influence the results found here.

REFERENCES

Borchers, D., Marques, T., Gunnlaugsson, T. and Jupp, P. 2010. Estimating distance sampling detection functions when distances are measured with errors. *J. Agric. Biol. Environ. Stat.* 15: 346–61.

Bøthun, G., Øien, N. and Skaug, H.J. 2008. Measurement error in survey and experiment, Norwegian minke whale surveys 2002–2007. Paper SC/60/PFI5 presented to the IWC Scientific Committee, June 2008, Santiago, Chile (unpublished). 8pp. [Paper available from the Office of this Journal].

Buckland, S.T., Anderson, D.R., Burnham, K.P. and Laake, J.L. 1993. *Distance Sampling: Estimating Abundance of Biological Populations*. Chapman and Hall, New York and London. xii+446pp.

International Whaling Commission. 2012. Requirements and Guidelines for Conducting Surveys and Analysing Data within the Revised Management Scheme. *J. Cetacean Res. Manage.* 13:509–17.

Marques, T.A. 2004. Predicting and correcting bias caused by error measurement in line transect sampling using multiplicative error models. *Biometrics* 60: 757–63.

Skaug, H.J., Øien, N., Schweder, T. and Bothun, G. 2004. Abundance of minke whales (*Balaenoptera acutorostrata*) in the northeastern Atlantic; variability in time and space. *Can. J. Fish. Aquat. Sci.* 61(6): 870–86.

Solvang, H.K., Skaug, H.J. and Øien, N.I. 2015. Abundance estimates of common minke whales in the Northeast Atlantic based on survey data collected over the period 2008–2013. Paper SC/66a/RMP08 presented to the IWC Scientific Committee, May 2015, San Diego, CA, USA (unpublished). 11pp. [Paper available from the Office of this Journal].

APPENDIX

A. According to eqn. (2) in Skaug *et al.* (2004) the bivariate probability density of the position of an initial sighting is given by

$$f(x,y) = w^{-1}\frac{\alpha}{v}Q(r_{x,y},\theta_{x,y})\exp\left\{-\frac{\alpha}{v}\int_y^\infty Q(r_{x,u},\theta_{x,u})\,du\right\},$$

where $r_{x,y} = (x^2 + y^2)^{1/2}$, $\theta_{x,y} = \arctan(x/y)$, V is the vessel speed, α is the surfacing rate, and $Q(r,\theta)$ is the hazard probability function.

Let $Q^{(1)}$ and $Q^{(2)}$ denote the hazard probability functions for the two platforms, 1 and 2, and denote by $u \in \{1,2,1{\cdot}2\}$ which platform made the initial observation ($1{\cdot}2$ indicates that the two platforms simultaneously observed the whale). Conditionally on the position of the initial sighting $(r,\theta) = (r_{x,y},\theta_{x,y})$, the trinomial probability distribution of u is given by

$$q(u) = \left\{Q^{1\cup 2}(r,\theta)\right\}^{-1}\cdot\begin{cases} Q^{(1)}(r,\theta)\left\{1-Q^{(2)}(r,\theta),\right\}, u=1 \\ Q^{(2)}(r,\theta)\left\{1-Q^{(1)}(r,\theta),\right\}, u=2 \\ Q^{(1)}(r,\theta)Q^{(2)}(r,\theta), \qquad u=1{\cdot}2 \end{cases}.$$

B. The effective strip half-width w is estimated by the following equation:

$$w = \int_0^\infty g(x)\,dx,$$

where $g(x)$ is the detection function (Buckland et al. 2001) defined as the probability that an animal located at perpendicular distance x from the transect line is detected. This equation corresponds to (1) in Skaug et al. (2004). Using the bivariate probability density $f(x,y)$ that introduced in Appendix, A, the detection function is given as

$$\begin{aligned} g(x) &= w\int f(x,y)\,dy \\ &= 1-\exp\left\{-\frac{\alpha}{v}\int_0^\infty Q(r_{x,y}\theta_{x,y})\,dy\right\} \end{aligned}$$

which corresponds to equation (3) in Skaug *et al.* (2004).

REFERENCE

Skaug, H.J., Øien, N., Schweder, T. and Bothun, G. 2004. Abundance of minke whales (*Balaenoptera acutorostrata*) in the northeastern Atlantic; variability in time and space. *Can. J. Fish. Aquat. Sci.* 61(6): 870–86.