

Effects of Beaufort Sea State on the Sightability of Harbor Porpoises in the Gulf of Maine

Debra Palka

National Marine Fisheries Service, Woods Hole, MA 02543, USA

ABSTRACT

Effects of Beaufort sea state heterogeneities on density estimates of harbor porpoises from data collected during 'independent' team line transect sighting surveys are investigated. The surveys were conducted in the Gulf of Maine/Bay of Fundy region during 1991 and 1992. It is shown that as Beaufort sea state increased (weather worsened), sighting rates (n/L) and the resulting density estimates markedly declined. At the same time, estimated probabilities of detecting the harbor porpoises [$f(0)$ and $\hat{g}(0)$] are reduced, but to a much lesser degree. Such results are expected if there is heterogeneity due to Beaufort sea state and the teams are conditionally independent. The relationship between Beaufort and density cannot be fully explained by small-scale spatial correlations between Beaufort sea state and estimated harbor porpoise density or by a team- or area-heterogeneity. The data could be interpreted that the Beaufort heterogeneity affected both the efficiency of the observers and the behavior of the harbor porpoises. Ways to compensate for the bias in this data set are discussed.

KEYWORDS: HARBOUR PORPOISE; ASSESSMENT; SURVEY-SHIP; $\hat{g}(0)$; BEHAVIOUR

INTRODUCTION

Using line transect sampling, the density of individuals, \hat{D} , is estimated from the observed sighting rates of groups, n/L , and the probability of detecting a group, $\hat{g}(0)/\hat{f}(0)$ using:

$$\hat{D} = \frac{n \cdot \hat{f}(0)}{2L \cdot \hat{g}(0)} \cdot \hat{E}(s) \quad (1)$$

where n is the observed number of groups seen, $\hat{f}(0)$ is the probability density function of observed perpendicular distances evaluated at distance zero, L is the length of the transect line, $\hat{g}(0)$ is the probability of detecting an animal group on the transect line, and $\hat{E}(s)$ is the expected value of group sizes, which, when there is no size-bias, is evaluated as the arithmetic mean of the observed group sizes (\bar{s}).

In 1991 and 1992, two line transect sighting surveys of harbor porpoises (*Phocoena phocoena*) in the Gulf of Maine/Bay of Fundy region were conducted using a sighting procedure where two 'independent' teams on the same ship searched the same body of water in front of the ship. It has been documented that when the two teams are searching the same area in front of the ship, they are actually conditionally independent. In which case, sources of heterogeneity can give rise to biases in the density estimate, in particular, the $\hat{g}(0)$ component (Schweder, 1990; Schweder *et al.*, 1991a; Buckland *et al.*, 1993). Possible sources of heterogeneity include: (1) observers (i.e. individuals within a team may have different sighting efficiencies); (2) platforms (i.e. platform characteristics may result in teams with different sighting probabilities); (3) animal behavior (e.g. animals displaying high visibility behaviors such as breaching may have a higher probability of been seen by both teams than an animal of the same species which is displaying a low visibility behavior); (4) environmental factors such as Beaufort sea state (i.e. physical characteristics of the ocean's surface or levels of sea sickness may cause sighting probabilities to differ between observers and/or for different individual animals); or (5) interactions between any of the above factors. Effects of heterogeneities on the density estimate can cause either a negative or positive bias, depending on the type of heterogeneity (Buckland *et al.*, 1993).

Prevailing weather conditions during a line transect survey appear to influence the number of animal groups seen. A reasonable expectation is that for a small species, such as harbor porpoises, individuals that could be detected while surveying in low Beaufort sea states (calm conditions, no white caps or swell) may be missed when surveying under high Beaufort sea states (rough conditions, white caps and/or swell). Thus, it is expected that sighting rates of harbor porpoises will decrease as Beaufort sea state increases. At the same time, as Beaufort sea state increases, it is expected that the probability of detecting a harbor porpoise will decrease; that is, $\hat{g}(0)$ decreases and/or $\hat{f}(0)$ increases. So, because density is a function of both sighting rate and probability of detection, it is expected that density estimated from data collected during different Beaufort sea states will be similar to each other. If heterogeneities exists, observed estimates of the parameters may not follow these expectations.

The sensitivity of $\hat{g}(0)$ to heterogeneities can be illustrated by the following example. Because the teams are conditionally independent, the sources of heterogeneities described above might give rise to a situation where highly visible animals are seen by both teams and cryptic animals are missed by both teams. In such cases, the number of sightings determined to have been seen by both teams (duplicate sightings) is biased upwards. As the degree of heterogeneity increases, the bias in number of duplicates also increases. If sighting functions for either the duplicates or teams do not change appreciably with changing conditions, then $\hat{g}(0)$ will be biased upwards, and consequentially, density is biased downwards.

This paper investigates the effect on the density estimate of heterogeneity due to Beaufort sea state. Two traditional methods of handling heterogeneities (stratification and covariance analyses) have been investigated. Sighting rates (n/L), $\hat{f}(0)$, $\hat{g}(0)$ and $\hat{E}(s)$ were estimated from data stratified by the Beaufort sea states experienced during the surveys (0 through 3) and then compared to each other. Density estimates were also compared when the sighting function was the half-normal or general-exponential covariate models, with Beaufort sea state being the covariate. To further explore the effects of Beaufort sea state, several other

potentially confounding factors were examined: (1) the degree of small-scale spatial correlation between Beaufort sea state and harbor porpoise density; (2) platform or team heterogeneities; (3) heterogeneities due to large-scale regions that contain different levels of expected density; and (4) animal behavior heterogeneity and its interaction with Beaufort sea state. Several ways to account for observed biases in this data set are discussed.

METHODS

Two line transect sighting surveys for harbor porpoises were conducted during 27 July to 30 August 1991 (Palka, 1995a) and 29 July to 3 September 1992 (Smith *et al.*, 1993). The study area was north of Portland, ME, USA, south of St. John's, New Brunswick, Canada and west of Port Joli, Nova Scotia, Canada (Fig. 1). The area was divided into two strata, offshore and inshore. The inshore stratum consisted of the shallow bays off Maine. The offshore stratum was made up of those waters outside the islands and shallow banks of Maine and Nova Scotia. In addition, the offshore stratum was further subdivided by the magnitude of expected harbor porpoise density: high density, intermediate density and low density. The analyses in this paper included only data collected within the offshore high density and intermediate density strata. This is because there were less than 10 sightings per year seen in the low density stratum, and the inshore stratum was surveyed with different ships in the two years.

The sighting platform, *R/V Abel-J*, data collection procedures, and half of the observers were the same during both years. The surveys were conducted when the Beaufort sea state was below 4, when visibility was greater than 500m and it was not raining. Data were collected independently by the two teams of observers, which were visually and acoustically separated. Both teams used the naked eye to

search for marine mammals. The teams were located in crow's nests on a mast at the bow of the ship, with unobstructed views of the water. The upper team, was 14m above the water line. The lower team, was directly below the upper team and at 9m above the water line. Each team consisted of four people, of which three were surveying and one was at rest. Every 30 minutes the members rotated from the port observation position, to the center observation position, to the starboard position, to rest, and then back to the port position. The port observer searched waters from 90° port to 10° starboard, where 0° was the trackline. The central observer searched from 30° port to 30° starboard. The starboard observer searched from 90° starboard to 10° port.

Sighting data collected included: location of each marine mammal group with respect to the ship; species identification; group size; direction the group was swimming; initial cue; and behavior of the group. Concurrently, the following environmental data were collected: Beaufort sea state; swell height and angle of approach; presence of fog and rain; cloud cover; direction and severity of glare; horizontal and vertical position of the sun; sea surface temperature; and wind direction and speed.

The Beaufort sea state scale was designed to determine the wind speed by using sea surface conditions (National Oceanic and Atmospheric Administration [NOAA], 1994). Beaufort sea state 0 describes a mirror-like sea surface that is due to the absence of wind (0 to 1 knots). Beaufort 1 indicates that wind speeds have increased to 1–3 knots and have produced a sea surface that has ripples on it that appear like scales, but there are no foam crests. Beaufort 2 indicates that winds have increased to 4–6 knots and have produced small, short wavelets that have a glassy appearance and do not break. Beaufort 3 indicates that wind speeds of 7–10 knots have produced large wavelets whose crests have begun to break, and so there are scattered white caps. Beaufort 4

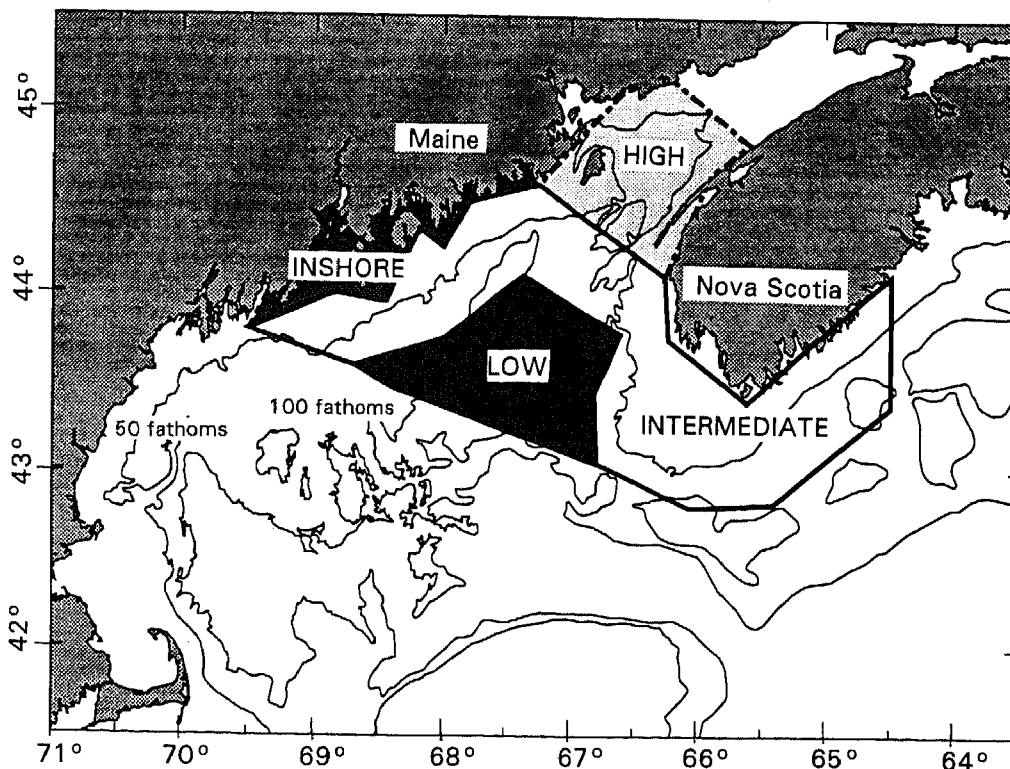


Fig. 1. Study area of line transect sighting surveys conducted during summers of 1991 and 1992. Shaded regions indicate the location of the high density, intermediate density, low density and inshore strata.

indicates that wind speeds of 11–16 knots have produced wavelets that are longer and there are fairly frequent whitecaps. The suggested wind speeds indicated above do not always produce the described sea surface conditions. This is particularly true in the Gulf of Maine/Bay of Fundy region where local effects of tides and interweaving water masses can at times influential the sea surface conditions more than the prevailing wind speeds. The tidal range within this area is from 8 to 26 feet (2.5 to 8m), and localized tidal currents can move at 2 to 5 knots (3.5–9km/hr). One example of localized effects is the presence of white caps; these may occur even when the wind is calm (less than five knots) if the wind is blowing in the opposite direction to a strong tidal current or alternatively be absent on other occasions even if the wind is blowing over 20 knots. For the present study, Beaufort refers to sea state conditions, rather than to wind speed.

To investigate effects of heterogeneities due to Beaufort sea state on various components of Equation 1, sightings data for each year were stratified by the Beaufort sea state experienced at the time of the sighting. Data from all predicted density strata were pooled to increase sample sizes. Later this condition was relaxed. Data from each Beaufort stratum were then used to estimate encounter rate (n/L), $esw = 1/\hat{f}(0)$, $\hat{g}(0)$, \bar{s} , and density, where esw is the effective strip width. The hazard rate model and the computer program DISTANCE (Buckland *et al.*, 1993) were used to estimate $\hat{g}(y)$. The maximum perpendicular distance (w) was 400m. The direct duplicate method was used to estimate $\hat{g}(0)$. Bootstrap re-sampling techniques were used to estimate confidence intervals for each parameter. A complete description of these methods are in Palka (1995a).

The estimate of $\hat{g}(0)$ for the upper team, $g_{up}(0)$, was:

$$\hat{g}_{up}(0) = \frac{n_{dup}}{n_{lo}} \cdot \frac{\int_0^w \hat{g}_{lo}(y) dy}{\int_0^w \hat{g}_{dup}(y) dy} \quad (2)$$

where n_{dup} is number of sightings seen by both teams, n_{lo} is number seen by the lower team, $\hat{g}_{lo}(y)$ is the detection probability of a sighting at perpendicular distance y as seen by the lower team, $\hat{g}_{dup}(y)$ is the detection probability of a sighting that was seen by both teams (duplicate sighting). A comparable equation was used to estimate $\hat{g}_{lo}(0)$.

Duplicate sightings were determined after the survey by two judges who independently reviewed the data and subjectively determined whether sightings were 'definite duplicates', 'possible duplicates' or 'unique' sightings. A definite duplicate was defined as a pair of sightings, one from each team, in which the time of detection, range and bearing were nearly exact, or it was highly likely that the animals traveled from the position seen first to the position seen second. A possible duplicate was defined as a pair of sightings in which it was less certain that they involved the same group of animals because, for instance, if measurement error was larger than usual then they could be the same group, or the animals were suspected to have started avoiding the ship and therefore, changed course or speed, or the group split or joined another group. All other sightings were categorized as unique sightings. It was assumed measurement error was greatest in the radial distance estimate, less in the bearing angle estimate and least in the time of detection. Duplicate sightings as determined by the two judges were similar. For example, for 1991, the ratio of total number of possible and definite duplicates to total

number of sightings detected by the upper team was 0.35 and 0.31 for judges A and B, respectively. The ratio for that seen in the high density stratum were slightly higher than the ratio for the intermediate density stratum (Judge A: 0.38 versus 0.31; Judge B: 0.33 versus 0.29). In general, relative to judge A, judge B assigned fewer possible duplicates and more definite duplicates. A complete comparison is given in Palka (1995a).

Two judges determining duplicate sightings results in two sets of duplicate sightings; there were thus two estimates of $\hat{g}(0)$ for each team. If it is assumed that these judges generated two random samples from the population of actual duplicates, one way to quantify the duplicate determination variability into the point estimate of $\hat{g}(0)$ and its confidence interval is to use both sets of duplicate sightings, following Schweder *et al.* (1991b). For each set of duplicate sightings, the original data were bootstrapped 1,000 times to produce 1,000 bootstrap estimates of $\hat{g}(0)$. Then, the 1,000 bootstrap estimates from each set of duplicates were combined into one joint distribution. The median of the joint bootstrap distribution was the point estimate of $\hat{g}(0)$, the standard deviation of the joint bootstrap distribution was the SE($\hat{g}(0)$), the 50th lowest value of the distribution was the lower 95% confidence limit, and the 50th largest value was the upper 95% confidence limit. These measures of variability were used in statistical tests to determine significant differences. The same bootstrap data was used to estimate $\hat{f}(0)$, n/L and \bar{s} .

The re-sampling unit in the bootstrap was a transect of survey effort within a density stratum, where a transect was defined as that section of the survey in which the ship was heading in the same direction during the same day. Within a bootstrap data set the number of transects in a stratum was constrained such that the length of track line within a stratum in the bootstrap data set equaled the length of track line actually surveyed within that stratum.

Three potentially confounding factors were also investigated. One was the underlying true small-scale harbor porpoise density patterns. For example, if areas with high Beaufort sea states happen to coincide with areas that are not favored habitats for harbor porpoises, then a relationship between Beaufort sea state and observed harbor porpoise density would be spurious. Although the true density patterns are not known, this potential relationship can be investigated by either determining the correlation between Beaufort sea state in regions that were not necessarily spatially grouped together and the associated harbor porpoise density for that region or by determining the correlation between Beaufort sea state and density in small spatial regions that presumably have similar true densities. To facilitate these inspections, for each year, a plot of the spatial distribution of Beaufort sea states was compared to the spatial distribution of densities. Both the Beaufort and density were measured for each 'leg of effort', where a leg of effort is defined as that continuous track line within a 'transect' in which all factors remain constant. This was usually 30 minutes or about 5 n.miles long. Density of harbor porpoises within a 'leg of effort' was estimated by the product of the sighting rate and average size of groups seen during that leg, and the $\hat{f}(0)$ and $\hat{g}(0)$ as estimated from data collected under the Beaufort sea state that was experienced during that leg (Palka, 1995b). These small scale density estimates were then kriged to create a contour map which was visually compared to the spatial map of Beaufort sea states. Simple kriging was used, where the variogram model was the spherical model, no drift or nugget was used, and anisotropy was at a 45° angle tilted NE-SW, which followed

the coastline. The computer package SURFER (Keckler, 1994) was used to create the kriged contour maps. Pearson's R^2 was used to quantify the correlation between Beaufort sea state and estimated density.

Other potential sources of heterogeneity include a team (or platform) effect and a large-scale expected density effect. The analyses described above used data pooled over (1) the entire study region which contains areas with different expected harbor porpoise densities and (2) both teams. To examine the influence of these two factors, density and its components (n/L , $\hat{f}(0)$, and \bar{s}) were estimated separately for each combination of team and expected density area. Because sample sizes within each team-area stratum were small, it was not reasonable to estimate $\hat{g}(0)$ for each stratum. Each team stayed on the same platform for the entire survey. Thus, it was not possible to distinguish between effects due to the individual observers within a team, and that due to the physical attributes of the platform that then influence the observers.

A third potential source of heterogeneity is the harbor porpoise's swimming behavior. It is possible that sightability of harbor porpoises depend on the animal's behavior pattern and that these behavior patterns may change when Beaufort sea state changes. To examine this factor, the distribution of behavior types was identified for each Beaufort sea state, and the associated sighting functions for each behavior were compared to each other. This was only done with data collected in 1992 because the definitions of behavior types were defined more precisely and the types of possible behaviors were expanded over that used in 1991. The behavior patterns were defined as low roll (1), high roll (2), repeated surfacing (3), porpoising (4), logging (5), splash-roll (6) and charging (7), where low rolls were the most inconspicuous behavior, and charging the most conspicuous. Definitions of each behavior type are given in Appendix A.

To describe the effects of heterogeneity, the above analyses stratified data by levels of the potential source of heterogeneity. Another way is to use covariate models that evaluate $\hat{g}(y)$. The covariate models investigated were the half-normal and general-exponential and the covariate was Beaufort sea state (Drummer and McDonald, 1987). Data collected in 1992 from each team were analyzed separately. The Beaufort covariate was defined in four ways. The first way was the usual integer values of Beaufort (e.g. 0, 1, 2 and 3). The second way was a 'real value' version of the above integer version (e.g. 0.2, 1.8, 2.4 and 3.9.) recorded while collecting the data (if this version is truncated then the usual integer definition of Beaufort results). The third definition used the exponential of definition 2 - this was introduced to attempt to capture the non-linearity arising out of the fact that only a small wind speed increase is needed to change from Beaufort 0 to 1 whereas a large increase is needed to change for Beaufort 2 to 3. The fourth approach used a weighted version of definition (3) as the sample size collected during Beaufort 3 was much smaller than for Beaufort 0. Since the log likelihood is maximized over the number of observations, the area of the surface which has more observations (Beaufort 0) is implicitly weighted more heavily, so that the fit is expected to be better there. To give the data collected in Beaufort 3 the same weight as that data collected in Beaufort 0, a simple weighted version of Beaufort was developed, where the number of exp(Beaufort) observations was inversely weighted to the relative density estimated for that Beaufort sea state. That is, the relative density for 1992 was approximately 1, 1, 2 and 8 for Beaufort sea states 0, 1, 2 and 3, respectively. So in this

version of the Beaufort covariate, for every Beaufort 3 observation seven more identical observations were added to the database.

RESULTS

Details of the results from the 1991 survey are described in Palka (1995a), and those from 1992 are in Smith *et al.* (1993). During both 1991 and 1992, approximately 71% of the sightings were detected while surveying in Beaufort sea states 0 and 1, which totalled about 35% of the track line length (Table 1). Sighting rates decreased as Beaufort sea state increased, although not linearly (Fig. 2). On average, sighting rates in Beaufort 1 were 0.8 of that in Beaufort 0, while sighting rates in Beaufort 2 and 3 were approximately 0.25 of that in Beaufort 0 (Fig. 3). Differences between sighting rates from Beaufort 0 and 3 were significant (z-test; $p < 0.01$). Sighting rates for the upper team were higher than that for the lower team during both years, although the overall differences were not significant ($p > 0.05$). Both the upper and lower team's sighting rates in 1992 were higher than their respective values in 1991, although not significantly ($p > 0.05$).

Table 1
Track line length and number of sightings detected by each team (upper and lower) during Beaufort sea states 0, 1, 2 and 3 during 1991 and 1992.

Sea state	Track length (% of total)		Number of sightings			
			Upper		Lower	
1991						
0	144	(8.4)	91	(0.63)	89	(0.63)
1	587	(34.1)	245	(0.42)	173	(0.29)
2	791	(46.1)	110	(0.14)	88	(0.11)
3	197	(11.5)	26	(0.13)	22	(0.11)
Total	1,720	(100.0)	472	(0.27)	372	(0.22)
1992						
0	150	(9.1)	85	(0.56)	95	(0.63)
1	589	(35.9)	350	(0.59)	324	(0.55)
2	661	(40.3)	158	(0.24)	112	(0.16)
3	240	(14.6)	22	(0.09)	21	(0.09)
Total	1,639	(100.0)	615	(0.38)	552	(0.34)

Estimates of effective strip width (esw) also decreased as Beaufort sea state increased (Fig. 2), although the rate of decline was less than that for sighting rates (Fig. 3). During Beaufort sea states of 0, 1 and 2, the esw was nearly the same; only in Beaufort 3 did the esw drop appreciably to 0.75 of that in Beaufort 0. Differences between the esw at Beaufort 0 and 3 were significant ($p < 0.01$). There were no consistent differences between years or between sighting teams. Perpendicular distances are estimated by radial distances between the ship and location of the harbor porpoise group and angles between the line of sight to the group and the line the ship is travelling on. The distributions of angles did not differ between sightings made under different Beaufort sea states. However, the distributions of radial distances followed the same pattern seen for esw , i.e. as Beaufort increased, the median radial distance decreased, with the largest decline between Beaufort 2 and 3.

Estimates of $\hat{g}(0)$; decreased when Beaufort sea state increased from 0 to 1 and from 1 to 2, however, at Beaufort 3, $\hat{g}(0)$; increased (Fig. 2). The maximum decrease of $\hat{g}(0)$ was at Beaufort 2, which was 0.65 of that from Beaufort 0 (Fig. 3); this difference was not significant ($p = 0.07$).

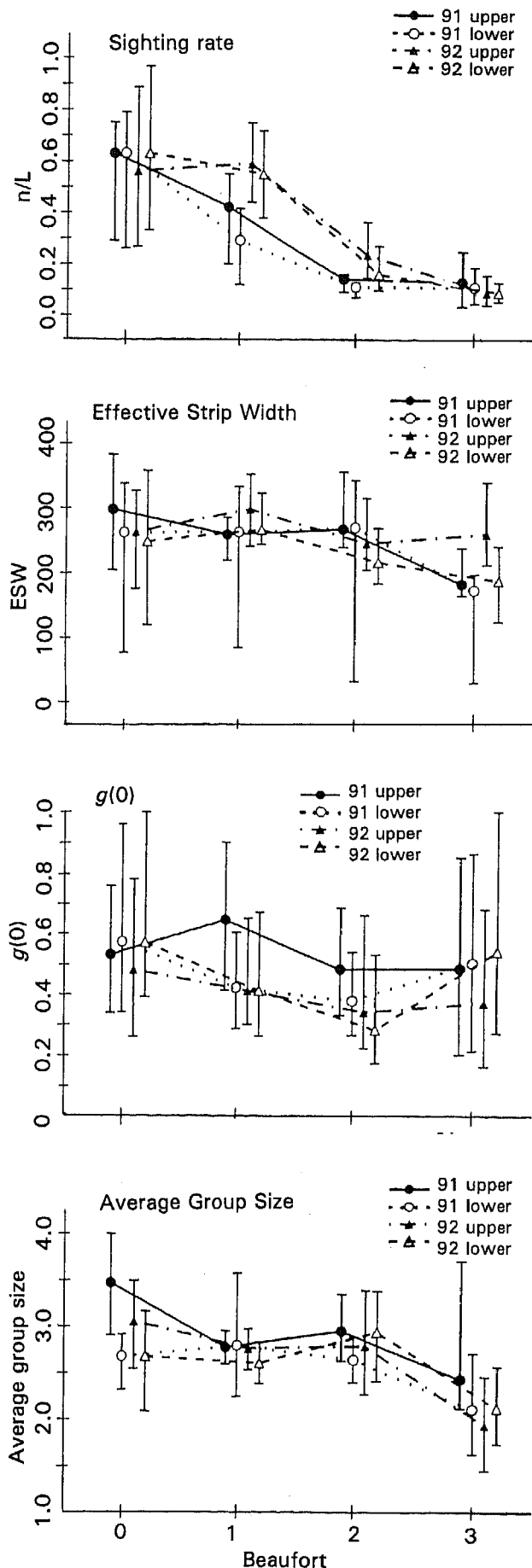


Fig. 2. Estimates of the sighting rates, effective strip width (*esw*), $\hat{g}(0)$ and average group size using data collected by the two platforms (up and low) during the Beaufort sea states experienced during the 1991 and 1992 surveys. Error bars are 95% confidence intervals.

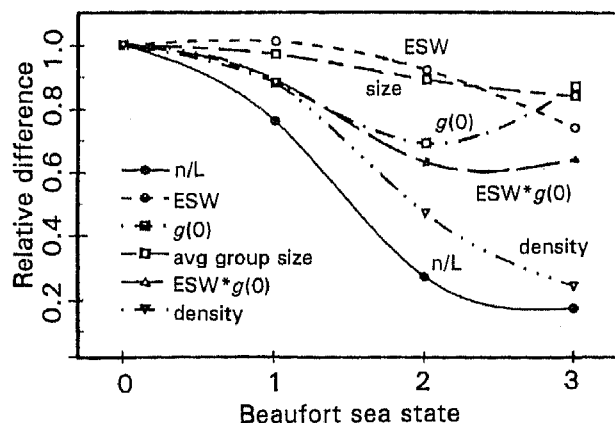


Fig. 3. Relative difference between Beaufort 0 and each subsequent Beaufort sea state for each of the following factors: sighting rates (*n/L*), effective strip width (*esw*), $\hat{g}(0)$, sightability [*esw** $\hat{g}(0)$], average group size (*size*), and density. A loess model was used to fit a line to the data points.

Average group size (\bar{s}) decreased slightly as Beaufort sea state increased (Fig. 2). Maximum difference was between Beaufort 0 and 3, where that from Beaufort 3 was 0.8 of that from Beaufort 0 (Fig. 3). This difference was significant ($p < 0.01$); however, the relative difference was much less than the relative differences in sighting rates and $\hat{g}(0)$.

The cumulative effects of the above factors is evident in the effect of Beaufort sea state on estimated density. As Beaufort sea state increased, estimated density decreased (Figs. 3 and 4). Density estimates were similar in Beaufort sea states 0 and 1, especially in 1992. As Beaufort increased to 2, density decreased by about half. As Beaufort increased to 3, in 1992 the density continued to decrease, while in 1991 the density leveled off. There was a significant difference between average density estimated for areas experiencing Beaufort 0 versus that from areas experiencing Beaufort 3 ($p < 0.01$).

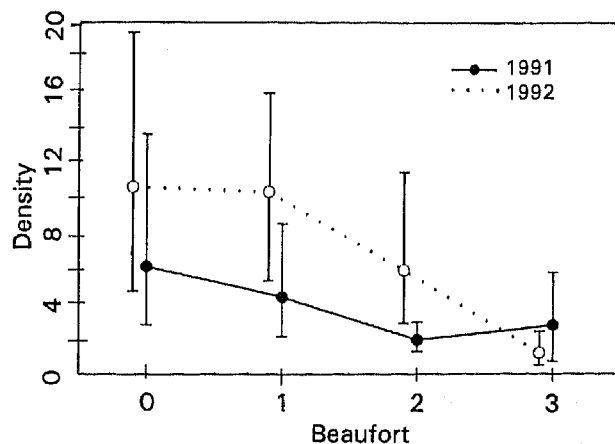


Fig. 4. For 1991 and 1992, the estimated density (animals/n.miles²) for areas surveyed under Beaufort sea states 0, 1, 2 and 3. Error bars are 95% confidence intervals.

For the entire survey area, when examining densities estimated from short lengths of trackline, i.e. on a small spatial scale, there was a significant correlation between estimated density of harbor porpoises and the Beaufort sea state experienced during that short length of track line (Pearson's correlation; 1991: $r = -0.15$ $p = 0.0011$; 1992: $r = -0.18$ $p = 0.0001$).

Areas where high Beaufort sea states were experienced were at times located close to areas that experienced low Beaufort sea states. At some of these places, for example, off the southern tip of Nova Scotia, where the harbor porpoise density is expected to be fairly consistent, during both 1991 and 1992, areas with high and low Beaufort sea states had similarly low densities. This resulted in a non-significant correlation between Beaufort and density (1991: $r=0.03$ $p=0.74$; 1992: $r=-0.08$ $p=0.38$). In other places such as the lower Bay of Fundy, legs with high and low Beauforts all had high densities and thus, there was a non-significant correlation between Beaufort and density (1991: $r=0.20$ $p=0.06$; 1992: $r=-0.03$ $p=0.74$). There were places such as off Penobscot Bay, Maine in 1991, where decreasing density contours coincided with increases in Beaufort; however, this also coincided with increasing depths and sea surface temperature (Palka, 1995b). The variability in densities within each Beaufort category indicate that areas which happen to be surveyed in high Beaufort sea states are not always areas with a true low harbor porpoise density.

Table 2

Sighting rates (n/L), effective strip width (ESW), average group size (\bar{s}) and density as estimated using data collected during 1992 by the upper and lower teams within different expected density strata (High, Intermediate and Total) and during different Beaufort sea states as experienced.

Stratum	Beaufort	n/L	ESW	\bar{s}	Density
Upper team					
High	0+1	0.90	410	2.96	60.0
	2	0.46	320	3.10	4.1
	3	0.41	311	2.53	3.1
Interm	0	0.69	352	2.91	5.2
	1	0.53	426	2.93	3.5
	2+3	0.09	222	2.70	10.0
Total	Stratified	0.40	319	2.85	2.9
	Not stratified	0.38	347	2.89	2.9
Lower team					
High	0+1	0.88	389	3.00	6.3
	2	0.40	255	3.32	4.8
	3	0.44	102	2.76	11.1
Interm	0	0.74	348	2.64	5.2
	1	0.53	337	2.39	3.4
	2+3	0.08	241	2.67	0.8
Total	Stratified	0.39	290	2.72	3.1
	Not stratified	0.37	324	2.71	2.9

When data were analyzed separately by year, team (or platform) and expected harbor porpoise density, the same patterns as described above occurred between Beaufort sea state and estimates of sighting rate, esw , \bar{s} , and density uncorrected for $\hat{g}(0)$ (Table 2). No consistent trends were observed between Beaufort sea state and the shape of the detection function for either team or in either the high density or intermediate density stratum (Palka, 1993). Sightings far from the trackline generally have a higher probability of being detected during Beaufort 0 than during Beaufort 3. However, at closer distances no pattern was apparent. In conclusion, there was no evidence that teams (or platforms) and areas with different expected densities were large sources of heterogeneity or helped explain the observed relationship between density and Beaufort sea state.

Table 3

Number of harbor porpoise groups that displayed the various swimming behaviors during Beaufort sea states 0, 1, 2 and 3. Definitions of behaviors are given in Appendix A. Percent of sightings in parentheses.

Behaviour	Beaufort Sea State			
	0	1	2	3
1 Low roll	255 (0.51)	451 (0.45)	70 (0.24)	8 (0.16)
2 High roll	50 (0.10)	81 (0.08)	41 (0.14)	5 (0.10)
3 Repeated	176 (0.35)	307 (0.31)	74 (0.26)	10 (0.20)
4 Porpoising	2 (0.00)	2 (0.00)	1 (0.00)	0 (0.00)
5 Logging	7 (0.01)	15 (0.01)	2 (0.01)	0 (0.00)
6 Splash-roll	4 (0.01)	42 (0.04)	41 (0.14)	4 (0.08)
7 Charging	6 (0.01)	102 (0.10)	57 (0.20)	24 (0.47)
8 Breaching	0 (0.00)	1 (0.00)	0 (0.00)	0 (0.00)
All	500	1,001	286	51

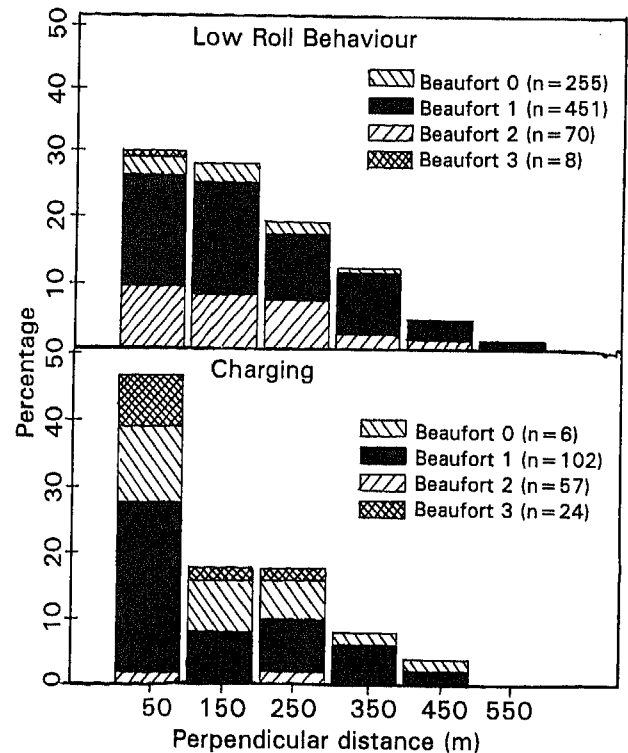


Fig. 5. Stacked histograms of the percent of sightings displaying either the 'low roll' behavior (top) or 'charging' behavior (bottom), where sightings were seen at the various perpendicular distances during Beaufort sea states 0, 1, 2 and 3. (n is the number of sightings seen at that Beaufort sea state)

Behaviors observed during Beaufort sea state 0 differed from that seen during Beaufort 3. During Beaufort 0, 51% of the harbor porpoises displayed a 'low roll' swimming behavior, and during Beaufort 3, only 16% were 'low rolling', i.e. displaying a low visibility behavior (Table 3). In contrast, 1.2% and 47% of sightings seen during Beaufort 0 and 3, respectively, were 'charging', a high visibility behavior. Sighting functions for these two behaviors differed (Fig. 5). Sighting functions for 'charging' animals were more peaked at distance 0, dropped off more sharply, and did not extend out as far as the sighting function for 'low rolling' animals. In addition, 'charging' animals were, in general, seen closer to the trackline than were 'low rolling' animals (Fig. 5). The distribution of radial distances for these two types of swimming behaviors indicate the same general patterns.

Drummer's covariate models of the half-normal and general exponential, where Beaufort sea state was a covariate, indicated that there was a significant effect of Beaufort, $\alpha < 0$ ($p < 0.05$; Table 4). The half-normal and general-exponential models fit the data well (χ^2 p -value < 0.05), particularly for the first three definitions of the Beaufort covariate. Estimated densities of groups were not very different when comparing estimates from the bivariate versus univariate model (Table 4). In all cases, the AIC (Akaike Information Criteria; Akaike, 1974) indicated that the bivariate model fitted better than the univariate model. This pattern held for all four definitions of the Beaufort covariate. The weighted exp(decimal Beaufort) covariate resulted in a higher density than that estimated with the other definitions of the Beaufort covariate.

Table 4

Results of an analysis of the 1992 data collected by the upper and lower teams where the covariate half-normal and covariate general exponential models were used to estimate the detection function. Results include: values of α (the effect of Beaufort sea state on the sighting function), p -value of the normal test where $H_0: \alpha = 0$ versus $H_a: \alpha < 0$, and the estimated density using the bivariate and univariate models.

Model	Covariate	α	p -value	Dens. bivar.	Dens. univar.
Upper team					
Half	Integer	-0.296	0.0005	1.03	1.01
Normal	Decimal	-0.301	0.0004	1.03	1.01
	Exp (Beauf)	-0.172	0.0001	1.04	1.01
	Weight Exp(B)	-0.176	0.0001	1.47	1.54
General	Integer	-0.278	0.0024	1.17	1.18
	Exp	-0.280	0.0023	1.17	1.18
Exp	Exp (Beauf)	-0.167	0.0001	1.17	1.18
	Weight Exp(B)	-0.183	0.0001	1.73	1.81
	Lower team				
Half	Integer	-0.190	0.0144	1.22	1.22
Normal	Decimal	-0.153	0.0312	1.23	1.22
	Exp (Beauf)	-0.086	0.0143	1.22	1.22
	Weight Exp(B)	-0.109	0.0346	1.91	1.89
General	Integer	-0.170	0.0346	1.35	1.39
	Exp	-0.162	0.0376	1.39	1.39
Exp	Exp (Beauf)	-0.070	0.0451	1.32	1.39
	Weight Exp(B)	-0.117	0.0000	2.32	2.31

DISCUSSION

If there were no heterogeneities due to Beaufort sea state, it is expected that the estimated density of harbor porpoises would be similar under different Beaufort sea states (as Beaufort increases, sighting rates may decrease, but e_{sw} and/or $\hat{g}(0)$; would compensate and therefore the product, density, would be nearly equivalent). Results of this study do not support this expectation. This study showed that as Beaufort sea state increased, fewer animals were detected (sighting rates decreased dramatically), but of those animals that were detected, their estimated detectability (e_{sw} and $\hat{g}(0)$) did not change very much, at least relative to the change in sighting rates. Similar results have been reported in Gunnlaugsson and Sigurjónsson (1990), Teilmann (1995) and Hammond *et al.* (1995).

Several other potentially confounding factors did not alter these results, and thus are probably of relatively minor influence. For example, there were no consistent correlations between small-scale spatial patterns of Beaufort sea state experienced and harbor porpoise density. Therefore, it is unlikely that *all* regions in which high Beaufort sea states were experienced are also regions with true low harbor

porpoises densities. However, in the waters off Penobscot Bay, Maine during 1991, as tracklines went offshore, sighting rates decreased, Beaufort sea state increased and, in addition, surface water temperatures increased to levels which were associated with low densities of harbor porpoises (Palka, 1995b). Thus, it is possible that in this particular region, the true harbor porpoise density decreased coincidentally as the Beaufort sea state increased.

Effects of heterogeneities due to team (or platform) and expected harbor porpoise density were minor. The same pattern between Beaufort sea state and estimated density occurred for each team and within each area of expected density.

Estimates of $\hat{g}(0)$ were shown to be biased upwards in increasing amounts as Beaufort sea state increased. As discussed in the introduction, the most likely explanation is a variation in availability of detecting a sighting (Schweder, 1990; Schweder *et al.*, 1991a; Buckland *et al.*, 1993). This effect is probably a result of heterogeneities due to Beaufort sea state. Although the mechanism is unclear, the present analysis indicated that for those times surveyed during Beaufort 2, the bias may be up to a factor of two and during Beaufort 3, up to a factor of three.

One interpretation of the data is that the changing Beaufort sea state affected both the efficiency of the observers and the behavior of the harbor porpoises. It is easy to see how changing physical properties of the ocean's surface could cause a small animal like a harbor porpoise more difficult for observers to see in higher Beaufort sea states. There are also indications that harbor porpoise swimming patterns might have changed as Beaufort increased, although behavioral heterogeneity need not be invoked to obtain the observed relationship between Beaufort sea state and density. The present data do not prove a behavioral change, but the rarity of 'charging' animals in low Beaufort sea states does perhaps suggest some change in behavior. In the absence of such a change, then the frequency of charging animals indicates that about one in ten animals were missed during Beaufort 1 compared to Beaufort 0, and one in fifty compared to Beaufort 3. These ratios are much higher than the sighting rates indicate. An alternative reason for the higher incidence of charging behavior in higher Beauforts is that 'high rolls' in low Beauforts cause splashes in high Beauforts. However, this seems unlikely because usually when splashes occur, more body is exposed than during the behavior defined as a high roll. Another hypothesis is that during high Beauforts, the higher ambient noise level in the ocean might mean that the ship is less detectable to the porpoise. It may then be startled by the ship when it is detected, and 'charge' away. There may be some truth in any of these scenarios at least some of the time. However, further work using radio-tagged animals is needed to definitively determine whether surfacing patterns change as Beaufort sea states change.

The two methods usually used to handle heterogeneities (stratification and covariate analyses) produced the same results, i.e. neither reduced the effects of the Beaufort heterogeneities. Other methods of handling heterogeneities include: (1) estimate density under low Beaufort sea states only (Hammond *et al.*, 1995); (2) use the variations in sighting rates with Beaufort sea states to provide a correction factor for the higher sea states (Gunnlaugsson and Sigurjónsson, 1990); (3) reduce the degree of dependence between teams on the same ship by having one team search farther ahead than the other team (Buckland *et al.*, 1993); and (4) reduce the degree of dependence even further by putting the two teams on different vessels (aerial survey in Hammond *et al.*, 1995). The first two methods could be done

retroactively with the present data. During a sighting survey conducted in 1995 in the Gulf of Maine/Bay of Fundy region, the third method above was attempted simultaneously with the sighting procedures described in this paper. Comparison of the two methods should be able to indicate if reducing the degree of dependence is sufficient to reduce the effects of Beaufort sea state heterogeneity.

Perhaps the major problem with any method involving two teams is the duplicate determination procedure: how to best determine which sightings were duplicates, how those duplicates are confounded with other factors such as underlying density or Beaufort sea states, and how sensitive is the density estimate to changes in the choice of duplicates. A comprehensive analysis of these topics is needed and considerable work is being undertaken in the context of obtaining abundance estimates for North Atlantic minke whales (e.g. see IWC, 1995; 1996).

The covariate method using the weighted version of Beaufort sea state presents another way to reduce effects of heterogeneities due to Beaufort sea state. This method may be of particular importance when 'correcting' data from surveys conducted in the past, where it is impossible to change the sighting procedures. This study indicated that density estimates did increase when the weighted Beaufort covariate definition was used. It is unknown whether this amount of increase is more representative of the true density. Simulations should be performed to investigate the effectiveness of this approach.

One consequence of the analysis presented here is that information on factors such as Beaufort sea state and swimming behaviors should be collected to determine the degree of potential heterogeneities, and if necessary be incorporated into the methods used to estimate absolute abundance, particularly for 'difficult-to-see' species, such as harbor porpoises. It is clear that the effect of including data collected in Beaufort sea states of greater than one, negatively bias an abundance estimate. However, the best way to account for this bias using such data needs to be investigated further.

ACKNOWLEDGEMENTS

I am indebted to the scientists and crew on the *R/V Abel-J* that help collect this data. In addition, I would like to thank Tim Smith and Gordon Waring from the National Marine Fisheries Service (166 Water St., Woods Hole, MA 02543) and two anonymous reviewers for comments on this manuscript.

REFERENCES

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on automatic control* 19(6):716-23.
- 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.
- Drummer, T.D. and McDonald, L.L. 1987. Size-bias in line transect sampling. *Biometrics* 43:13-21.
- Gunnlaugsson, T. and Sigurjónsson, J. 1990. NASS-87: Estimation of whale abundance based on observations made onboard Icelandic and Faroese survey vessels. *Rep. int. Whal. Commn* 40:571-80.
- Hammond, P.S., Benke, H., Berggren, P., Borchers, D.L., Buckland, S.T., Collet, A., Heide-Jørgensen, M.P., Heimlich-Boran, S., Hiby, A.R., Leopold, M.F. and Øien, N. 1995. Distribution and abundance of the harbour porpoise and other small cetaceans in the North Sea and adjacent waters. Final Report to the EU Commission. LIFE 92-2/UK/O27.
- International Whaling Commission. 1995. Report of the Scientific Committee, Annex H. Report of the Working Group on abundance estimation. *Rep. int. Whal. Commn* 45:187-94.
- International Whaling Commission. 1996. Report of the Scientific Committee, Annex I. Report of the Working Group on abundance estimation. *Rep. int. Whal. Commn* 46
- Keckler, D. 1994. *SURFER for Windows User Guide*. Golden Software Inc, CO, USA. 483pp.
- National Oceanic and Atmospheric Administration [NOAA]. 1994. *United States Coast Pilot*, Washington DC. 170pp.
- Palka, D. 1993. Estimating density of animals when assumptions of line-transect surveys are violated. Ph.D. Thesis, Scripps Inst. of Oceanography, University of California, San Diego. 169pp.
- Palka, D. 1995a. Abundance estimate of the Gulf of Maine harbor porpoise. *Rep. int. Whal. Commn* (special issue 16):27-50.
- Palka, D. 1995b. Influences on spatial patterns of Gulf of Maine harbor porpoises. pp. 69-75. In: A.S. Blix, L. Walløe and Ø. Ulltang (eds.) *Developments in Marine Biology*. 4. *Whales, Seals, Fish and Man: proceedings of the International Symposium on the Biology of Marine Mammals in the Northeast Atlantic*. Tromsø, Norway, 29 November-1 December 1994. Elsevier Science B.V., The Netherlands. 720pp.
- Schweder, T. 1990. Independent observer experiments to estimate the detection function in line transect surveys of whales. *Rep. int. Whal. Commn* 40:349-55.
- Schweder, T., Host, G. and Øien, N. 1991a. A note on the bias in capture-recapture type estimates of $g(0)$; due to the fact that whales are diving. *Rep. int. Whal. Commn* 41:397-9.
- Schweder, T., Øien, N. and Host, G. 1991b. Estimates of the detection probability for shipboard surveys of northeastern Atlantic minke whales, based on a parallel ship experiment. *Rep. int. Whal. Commn* 41:417-32.
- Smith, T.D., Palka, D. and Bisack, K. 1993. Biological significance of bycatch of harbour porpoise in the Gulf of Maine demersal gillnet fishery. NOAA/NMFS/NEFSC, Woods Hole, MA. Northeast Fisheries Science Center, Reference Document 93-23. 15pp.
- Teilmann, J. 1995. Influence of seastate on abundance estimates of harbour porpoise (*Phocoena phocoena*). Paper SC/47/SM45 presented to the IWC Scientific Committee, May 1995 (unpublished). 27pp.

Appendix A

DESCRIPTION OF BEHAVIOR CODES

- | | | |
|---|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Low roll | almost invisible, fin and perhaps the top of the back only are seen, it may seem that the animal is trying to avoid being seen. |
| 2 | High roll | more obvious roll, with the flanks of the animal visible in a distinctive arch, it appears the animal may have surfaced oblivious of observers or may be preparing to dive deep. |
| 3 | Repeated surfacing | the animal is observed several times surfacing 'normally', each surfacing at some irregular distance from the previous one; it appears to be just moving along. |
| 4 | Porpoising | the animal is rolling along slowly at the surface, breaking the surface regularly every few body lengths. |
| 5 | Logging | the animal appears to be sleeping. |
| 6 | Splash-roll | the animal rolls high out of the water creating splashes, often repeating this behavior; it appears to be feeding, fighting, or mating when more than one is present. |
| 7 | Charging | the animal is charging through waves moving very fast and creating a lot of splashes. |